POLYPLOIDY AND EVOLUTION OF BREEDING SYSTEMS IN AN ARCTIC-ALPINE SPECIES COMPLEX OF *PRIMULA*: AN INTEGRATIVE PHYLOGENETIC AND CYTOGENETIC STUDY

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftliche Fakultät

der

Universität Zürich

von

Alessia Guggisberg

von

Zimmerwald/BE

Promotionskomitee

Prof. Dr. Elena Conti (Vorsitz)
Prof. Dr. Jakob Schneller
Prof. Dr. Alex Widmer

Zürich, 2008

An meine Eltern

Table of Contents

Summary

Polyploidy (the presence of more than two chromosome sets in a nucleus) is a major mechanism of adaptation and speciation in plants. Two types of polyploids are commonly recognised. *Auto*polyploids form within a single species and contain more than two sets of homologous chromosomes in their nuclear genome, whereas *allo*polyploids contain more than two sets of homeologous chromosomes that diverged from each other prior to interspecific hybridisation. It has been early hypothesised that polyploid plants evolve higher rates of self-fertilisation than their diploid relatives. Furthermore, the frequency of polyploids has been repeatedly shown to increase with latitude. Altogether, these two observations suggest an association between polyploidy, self-compatibility and high-latitudinal distribution. *Primula* sect. *Aleuritia* subsect. *Aleuritia* (Primulaceae – hereafter *Aleuritia*) constitutes an ideal model system to investigate this proposed correlation. *Aleuritia* consists of 21 arctic-alpine species, which display variation of ploidy levels, from diploid to tetradecaploid, and breeding systems, alternating between self-incompatible heterostyly to self-compatible homostyly. Indeed, all eight diploid species are heterostylous, whereas ten of the 11 polyploid species are homostylous. Furthermore, the highest polyploids do not occur south of latitude 45°N or north of latitude 40°S, whereas the diploids usually occur between these latitudes.

The main goal of the present thesis was to disentangle the reticulation patterns in *Aleuritia*, and to investigate the consequences of polyploidy on the evolution of breeding systems, using an integrative and multifaceted approach. In **Chapter 2**, we elaborate a maternal phylogeny based on chloroplast (cpDNA) sequences to gain insight into the biogeographic history of the group, and explore the switches in breeding system, ploidy level and distributional range, to work out an explanation for the success of polyploid taxa in extreme arctic environments. The cpDNA phylogeny confirmed the monophyly of *Aleuritia* and suggested that an Asian ancestor gave origin to an amphi-Pacific clade and to a lineage that diversified on the European and American continents, respectively. The phylogenetic analyses inferred a diploid, heterostylous most recent common ancestor for *Aleuritia* and multiple origins for the polyploid, homostylous lineages. They further confirmed the proposed association between polyploidy and homostyly. Comparisons of distributional ranges between homostylous and heterostylous species finally showed that autogamous taxa expanded further north into previously glaciated areas than allogamous ones, implying that the switch to self-compatibility may have represented a selective advantage in the colonisation of empty ecological niches freed up after glacier retreats. Overall, the high frequency of polyploid,

autogamous species in the Arctic may be the product of selection for increased selfing ability in habitats where pollination is unreliable.

In **Chapter 3**, we generate direct and cloned nuclear (nDNA) sequences to unravel the reticulation patterns caused by allopolyploidy. These investigations indicated that the South-American octoploid *Primula magellanica* originated within the amphi-Pacific lineage, and advocated for repeated cycles of reticulation between the amphi-Pacific, the European and the North-American lineages for most other polyploids within *Aleuritia*. Indeed, the use of different types of data and multiple analytical approaches in combination with pre-existing knowledge from morphological, caryological and distributional data, allowed us (i) to infer an allopolyploid origin for most polyploids within *Aleuritia,* and (ii) to propose a recurrent origin for at least two of them.

In **Chapter 4**, we develop *in situ* hybridisation (ISH) techniques to verify the hybrid origin of an intersectional allopolyploid, *Primula egaliksensis*, and to assess the extent and localisation of genomic reorganisation within the parental genomes. Results from ISH experiments confirmed the hybrid origin of *P. egaliksensis* and advocated for a "geologically young" age of this taxon. They further revealed no large-scale, but drastic small-scale genomic rearrangements targeted at the maternal genome. The lack of major intergenomic rearrangements may be a consequence of the considerable genetic divergence between the putative parents, while the minor intragenomic rearrangements may be indicative of first steps towards genome diploidisation.

Altogether, this study is compatible with the secondary contact model earlier invoked to explain speciation in *Aleuritia*. According to this model, glacial advancement during the Pleistocene caused the fragmentation of diploid, allogamous populations, which may have survived in ice-free areas (i.e. refugia). As glaciers retreated, the differentiated diploid populations came into contact again and hybridised, giving origin to autogamous, allopolyploid taxa. The establishment and further success of these newly formed polyploids may be attributed to the switch from heterostyly to homostyly, i.e. the switch from obligate outcrossing to self-compatibility.

In **Chapter 5**, I enumerate some suggestions for future investigations. Because most polyploids of *Aleuritia* are of hybrid origin, it is still unclear whether hybridisation or polyploidisation was the driving force for the evolution of homostyly. Similarly, it remains to be clarified whether recombination constitutes the genetic mechanism causing the switch in breeding system. Lastly, the incorporation of a temporal framework would allow one to test whether the Quaternary glaciations actually played a central role in the evolution of *Aleuritia*.

Zusammenfassung

Polyploidie (das Vorhandensein von mehr als zwei Chromosomensätzen pro Zellkern) ist einer der Hauptmechanismen der Anpassung und Artbildung bei Pflanzen. Zwei Polyploidie-Typen werden gewöhnlich unterschieden. *Auto*polyploidie bildet sich innerhalb einer Art und besteht aus mehr als zwei homologen Chromosomensätzen im Zellkern, während *Allo*polyploidie aus dem Zusammenfügen von mehr als zwei unterschiedlichen Chromosomensätzen bei zwischenartlicher Hybridisierung resultiert. Gemäss einer schon lange bekannten Hypothese neigen polyploide Pflanzen mehr zu Selbstbestäubung als ihre diploiden Artgenossen. Mehrfach wurde auch gezeigt, dass die Häufigkeit von Polyploiden mit der geographischen Breite zunimmt. Zusammen deuten diese Beobachtungen auf ein vermehrtes Auftreten von Polyploidie und Selbstkompatibilität in hohen Breiten. Um diese Wechselbeziehung zu erforschen, bildet die Artengruppe von *Primula* sect. *Aleuritia* subsect. *Aleuritia* (Primelgewächse – zukünftig *Aleuritia* genannt) ein ideales Modellsystem. *Aleuritia* besteht aus 21 arktischen und alpinen Arten, deren Ploidiegrad zwischen Diploidie und Tetradecaploidie schwankt, und die zwischen selbstinkompatibler Heterostylie und selbstkompatibler Homostylie abwechseln. In der Tat sind alle acht diploiden Arten heterostyl, während zehn der insgesamt 11 polyploiden Arten homostyl sind. Ausserdem kommen die *Aleuritia*-Arten mit höheren Polyploidiestufen zwischen 45° nördlicher Breite und 40° südlicher Breite selten vor, während die Diploiden hauptsächlich zwischen diesen Breitengraden gedeihen.

Das Hauptziel der vorliegenden Dissertation war herauszufinden, wie die Arten der *Aleuritia*-Gruppe miteinander verwandt sind und welche Konsequenzen das Auftreten von Polyploidie auf die Reproduktionsbiologie der einzelnen Arten hatte. In **Kapitel 2** wird eine auf Chloroplasten-Sequenzen (cpDNA) basierende mütterliche Phylogenie vorgestellt, um einen Einblick in die Biogeographie der Gruppe zu geben und eine Hypothese darüber aufzustellen, weshalb Polyploide in extrem arktischen Biotopen eine besonders hohe Erfolgsrate vorweisen. Die auf cpDNA basierende Phylogenie bestätigte die Monophylie von *Aleuritia* und deutete darauf hin, dass sich eine amphi-pazifische Linie aus einem asiatischen Vorfahren ableitete. Daraus entstand auch eine Linie, die in Europa und Amerika neue Arten entwickelte. Den phylogenetischen Analysen folgend war der jüngste Vorfahre von *Aleuritia* diploid und heterostyl, wobei daraus mehrfach polyploide und homostyle Arten entstanden sind. Ausserdem konnte die oben erwähnte Korrelation von Polyploidie und Homostylie bestätigt werden. Schliesslich zeigten Vergleiche von Verbreitungsarealen homostyler und

heterostyler Arten, dass selbstkompatible Arten in früher vereisten Regionen weiter nach Norden vordringen konnten als ihre selbstinkompatiblen Artgenossen. Die Fähigkeit zur Selbstbestäubung scheint demnach ein Selektionsvorteil bei der Kolonisierung frei werdender Gebiete nach dem Rückzug der Gletscher gewesen zu sein. Selbstkompatible Polyploiden profitierten davon von der Möglichkeit, sich in Habitaten zu etablieren, in denen die Fremdbestäubung unzuverlässiger war als Selbstbestäubung.

Im **Kapitel 3** werden direkte und klonierte nukleare Sequenzen genutzt, um die durch Allopolyploidie verursachten Polyploidkomplexe zu rekonstruieren. Diese Studie zeigte, dass die in Südamerika vorkommende octoploide Art *Primula magellanica* von einer amphi-pazifischen Linie abstammt, und dass die meisten übrigen Polyploiden aus wiederholten Hybridisierungszyklen zwischen amphi-pazifischen, europäischen und nordamerikanischen Linien resultieren. Zusätzlich erlaubte uns der Einsatz verschiedener molekularer Datensätze und Erkenntnisse zur Morphologie, Karyologie und Biogeographie der Arten: (i) eine allopolyploide Herkunft für die meisten Polyploiden in *Aleuritia* vorzuschlagen, und (ii) eine mehrfache Entstehung von mindestens zwei polyploiden Arten zu postulieren.

Das **Kapitel 4** behandelt *in situ* Hybridisierungstechniken, mit deren Hilfe die Herkunft des allopolyploiden Bastards *Primula egaliksensis* überprüft und genomische Restrukturierungen der aus verschiedenen Sektionen stammenden elterlichen Genome aufgedeckt werden können. Die Resultate der Hybridisierungsexperimente bestätigten die hybride Herkunft von *P. egaliksensis* und deuteten auf ein junges „geologisches Alter“ dieser Art hin. Die genomischen Restrukturierungen im mütterlichen Genom sind nicht grossflächig, dafür aber drastisch und kleinflächig. Das Fehlen von deutlichen intergenomischen Restrukturierungen könnte die Konsequenz beträchtlicher genetischer Differenzen zwischen beiden Eltern sein, während die geringeren intragenomischen Restrukturierungen wohl erste Anzeichen für die Diploidisierung des Genoms darstellen.

Zusammenfassend darf geschlossen werden, dass die vorliegenden Untersuchungen das „secondary contact model“ bestätigt, welches schon früher zur Erklärung der Artbildungsprozesse von *Aleuritia* vorgeschlagen worden ist. Diesem Modell entsprechend haben Gletscher während der Eiszeit das Zersplittern diploider selbstinkompatibler Populationen gefördert, die anschliessend Zuflucht in eisfreien Gebieten (sog. Refugien) fanden. Als sich die Eismassen zurückzogen, kamen diese differenzierten diploiden Populationen wieder in Kontakt miteinander und hybridisierten, woraus allopolyploide selbstkompatible Arten resultierten. Einiges spricht dafür, dass Etablierung und Erfolg dieser

neu gebildeten Polyploiden vom Wechsel von Heterostylie zur Homostylie abhingen, d. h. vom Wechsel von obligater Fremdbestäubung zur Selbstkompatibilität.

In **Kapitel 5** werden Vorschläge für zukünftige Projekte aufgelistet. Da bei *Aleuritia* die meisten Polyploiden eine hybride Herkunft haben, ist es noch unklar, ob Hybridisierung oder Polyploidisierung das Auftreten der Homostylie bewirkt haben. Gleichermassen unklar bleibt, ob die Rekombination selber den genetischen Mechanismus darstellt, der den Wechsel zur Homostylie verursacht. Schliesslich würden molekulare Datierungen ermöglichen abzuschätzen, ob die Quartäreiszeiten eine wirklich zentrale Rolle in der Entwicklung von *Aleuritia* gespielt haben.

Chapter 1

Introduction and thesis outline

Polyploidy in plant evolution

Polyploidy (i.e. the presence of more than two chromosome sets in a nucleus) is a major mechanism of adaptation and speciation in plants (Soltis and Soltis, 1993; Bretagnolle *et al.*, 1998; Ramsey and Schemske, 1998; Levin, 2002; Soltis *et al.*, 2003). Otto and Whitton (p. 427, 2000) even stated that 'polyploidisation may be the single most common mechanism of sympatric speciation in plants' (see Box 1). Indeed, up to 70% of angiosperms have been proposed to be of polyploid origin (Stebbins, 1950; Grant, 1981; Masterson, 1994; Otto and Whitton, 2000), including species of small genome size and chromosome number such as *Arabidopsis thaliana* (Vision *et al.*, 2000; Henry *et al.*, 2006).

Two types of polyploids are commonly recognised, according to the degree of homology among co-existing genomes (Ramsey and Schemske, 1998). *Auto*polyploids form within a single species and contain more than two sets of homologous chromosomes in their nuclear genome, while *allo*polyploids contain more than two sets of homeologous chromosomes that diverged from each other prior to interspecific hybridisation (Fig. 1). Hence, autopolyploids are generally characterised by the formation of multivalents at meiosis, and allopolyploids by bivalent pairing (Fig. 1). Allopolyploids are thought to be the predominant form of polyploidy in flowering plants, although this hypothesis has not been fully demonstrated (Soltis *et al.*, 2003). Importantly, hybridisation may also give origin to homoploid hybrids that maintain the same ploidy level as their parents (Arnold, 1997; Rieseberg and Carney, 1998; Arnold, 2006).

Box 1. Mechanisms of speciation*

Speciation is usually seen as a branching process (**cladogenesis**) in which new kinds of organisms originate from a single ancestral species, but new species may also arise through hybridisation (i.e. **reticulation**).

Allopatric speciation may occur when a species is subdivided by a barrier to dispersal (e.g. rising of sea level or mountain chain) into two populations which later evolve reproductive isolation as a result of their geographical isolation.

Sympatric and **parapatric speciation** define the divergence of spatially largely- and narrowly-overlapping populations, respectively. Two isolating mechanisms have been proposed: (i) disruptive selection, in which selective pressures cause a population to adapt to a different niche (e.g. switch in host niche); and (ii) chromosomal changes (e.g. polyploidy).

*after Lomolino *et al.* (2006)

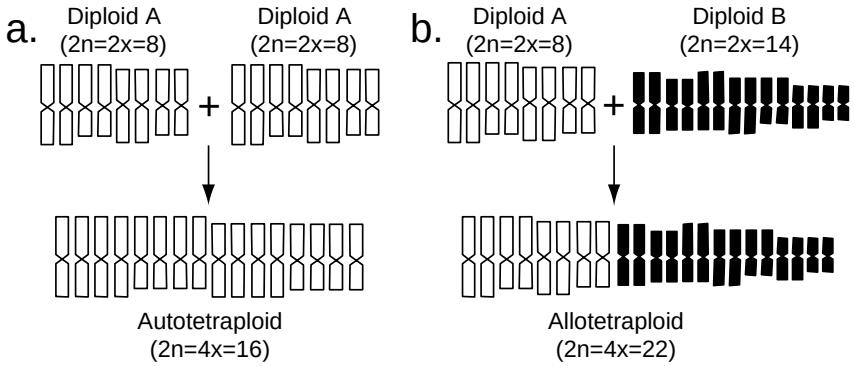


Figure 1. Origin of an autopolyploid (a) and an allopolyploid (b). Redrawn from Leitch and Bennett (1997).

An abundant literature has been devoted to the mode of polyploid formation (Harlan and DeWet, 1975; Thompson and Lumaret, 1992; Bretagnolle and Thompson, 1995; Ramsey and Schemske, 1998; Bretagnolle, 2001). Although in some taxa somatic doubling may be responsible, fertilisation between unreduced gametes is thought to be the predominant mode of polyploid emergence (Fig. 2). Furthermore, recent analysis of polyploid genomes has considerably strengthened the idea that multiple origins are common in generating polyploids (Soltis and Soltis, 1993, 1999; Soltis *et al.*, 2003). Few examples of polyploid taxa that seem to have originated once have been documented (e.g. *Spartina anglica*, Ainouche *et al.*, 2004).

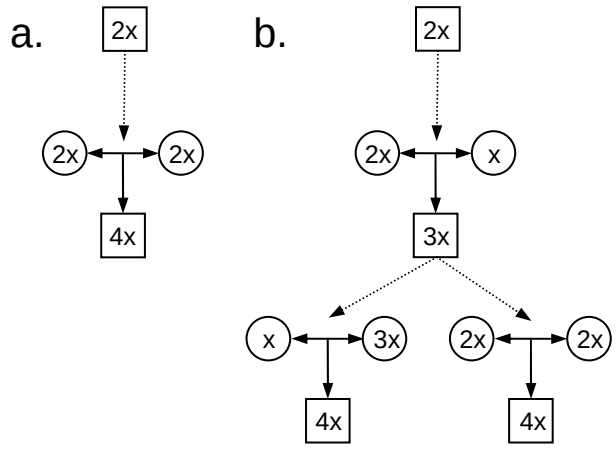


Figure 2. Sexual mechanisms of polyploidisation. a) Bilateral polyploidisation involves the fusion of two unreduced gametes. b) Unilateral polyploidisation involves the fusion of an unreduced gamete with a reduced gamete to produce a triploid. Triploids can subsequently generate tetraploids by backcrossing to diploids, self-fertilising, or by crossing with each other (triploid bridge hypothesis). Gametophytes are represented by squares, and gametes by circles. According to Harlan and deWet (1975) and Bretagnolle and Thompson (1995).

When polyploids arise within diploid populations, they form a very small minority (minority-cytotype disadvantage) and risk to become extinct (minority cytotype exclusion principle; Levin, 1975), because they are mainly fertilised by haploid pollen and produce highly sterile triploids (triploid block effect). This disadvantage may be overcome under following conditions: (i) competitive advantage of the polyploids (Fowler and Levin, 1984; Felber, 1991; Rausch and Morgan, 2005); (ii) high frequency of diplogametes (Felber, 1991; Rausch and Morgan, 2005); (iii) small diploid population (Fowler and Levin, 1984; Rausch and Morgan, 2005); (iv) shift in ecological niches (Fowler and Levin, 1984; Rodríguez, 1996; Husband and Sabara, 2003); (v) assortative mating (i.e. asynchronous flowering or switch in pollinator preferences; Segraves and Thompson, 1999; Husband and Sabara, 2003; Rausch and Morgan, 2005); (vi) self-compatibility (Rodríguez, 1996; Husband and Sabara, 2003; Rausch and Morgan, 2005); and (vii) apomixis (Levin, 2002).

Polyploids are thought to undergo rapid genome restructuring following their formation (Wendel, 2000; Adams and Wendel, 2005; Chen and Ni, 2006) as evidenced, for instance, by the studies on synthetic and natural allopolyploids of *Brassica* (Song *et al.*, 1995), *Arabidopsis* (Pontes *et al.*, 2004; Wang *et al.*, 2004) and *Nicotiana* (Skalická *et al.*, 2003; Kovarík *et al.*, 2004; Lim *et al.*, 2004; Skalická *et al.*, 2005). Those changes range from intergenomic chromosome translocations to interlocus recombinations (e.g. unequal crossing-over and gene conversion), which can be accompanied by epigenetic modifications at some loci (such as cytosine methylation) leading to modifications in gene expression patterns (Wendel, 2000; Liu and Wendel, 2003; Chen and Ni, 2006). Altogether those alterations may constitute a response to the stress imposed by hybridisation and/or chromosome doubling (i.e. genomic shock; McClintock, 1984; Madlung and Comai, 2004), and result in the gradual diploidisation of polyploid taxa (Leitch and Bennett, 2004; Ma and Gustafson, 2005). Yet, there are also examples of polyploids which underwent only few changes in overall genome structure since their formation, e.g. *Gossypium* (Liu *et al.*, 2001) and *Spartina* (Baumel *et al.*, 2002; Ainouche *et al.*, 2004).

Correlation between polyploidy, self-compatibility and latitude

It has been early hypothesised that polyploid plants evolve higher rates of self-fertilisation than their diploid relatives (Stebbins, 1950; Grant, 1956; Barringer, 2007; Husband *et al.*, 2008). Three reasons (which are not mutually exclusive) have been invoked for the apparent correlation between polyploidy and selfing rate. First, polyploidy may cause the breakdown of self-incompatibility systems (see Box 2), enhancing the probability of self-fertilisation (Barrett, 1988; Richards, 1997). Second, newly-arisen polyploids (i.e. *neo*polyploids) may be more likely to establish if self-compatible, because they would avoid wasting their pollen in sterile triploid offspring, i.e. reduce

the risk of extinction via negative frequency-dependent selection (cf. minority cytotype exclusion principle; Levin, 1975; Rodríguez, 1996; Ramsey and Schemske, 1998; Rausch and Morgan, 2005). Finally, polyploids (especially neopolyploids) may be buffered against the detrimental effects of inbreeding depression (lower fitness of inbred lines compared to outbred lines), due to the presence of multiple alleles and the associated reduction in homozygosity, resulting in increased level of self-fertilisation (Lande and Schemske, 1985; Husband and Schemske, 1997).

Box 2. Types of self-incompatibility systems*

Self-incompatibility systems are mechanisms that promote outcrossing, because no viable seeds can be produced upon self-pollination. They can be divided into two categories:

a) **gametophytic self-incompatibility**, in which the incompatibility phenotype of the pollen is determined by its own haploid genotype;

b) **sporophytic self-incompatibility**, in which the incompatibility phenotype is determined by the genotype of the pollen-producing parent.

Whereas all gametophytic self-incompatibility systems consist of only one floral morph (i.e. are homomorphic), sporophytic self-incompatibility systems can be further subdivided into homomorphic and heteromorphic systems depending whether one or more floral morphs are present in the populations.

*after Barrett (1988) and Richards (1997)

The frequency of polyploids has been repeatedly shown to increase with latitude, especially the proportion of polyploids above the tetraploid level (e.g. Löve and Löve, 1957; Stebbins, 1971; Brochmann *et al.*, 2004). It was first suggested that polyploids are better adapted to extreme cold than their diploid congeners (increased hardiness hypothesis; Stebbins, 1950), but this assumption was later refuted on the basis of following observations: (i) the frequency of polyploidy does not differ between the Alps and surrounding lowlands, and (ii) polyploids are more common among perennial herbs, which are the dominant growth form in very cold regions (Stebbins, 1971; Levin and Wilson, 1976).

The distribution and composition of the arctic flora was probably greatly affected by the climatic oscillations of the Quaternary, which caused the retreat of arctic plants in ice-free areas (refugia) during glacial advancements and their subsequent expansion during interglacial periods (reviewed in Abbott and Brochmann, 2003). Consequently, it has been hypothesised that polyploids are more frequent at higher latitudes, because they are more successful at recolonising newly deglaciated (i.e. disturbed) areas than diploids (Stebbins, 1971; Ehrendorfer, 1980). Stebbins (1984) refined this assumption and postulated a positive correlation between the frequency of polyploids and the degree of glaciation (rather than latitude), as recently demonstrated by Brochmann *et al.* (2004). Several reasons have been proposed to explain the evolutionary success of polyploids at recolonising previously glaciated areas: (i) high levels of heterozygosity (Brochmann and Elven, 1992; Thompson and Lumaret, 1992; Soltis and Soltis, 2000; Brochmann *et al.*, 2004), (ii) genetic

novelty (e.g. resulting from transgressive segregation; Rieseberg *et al.*, 1999), and (iii) novel nucleo-cytoplasmic interactions (Levin, 2002).

The genus *Primula* as a model system

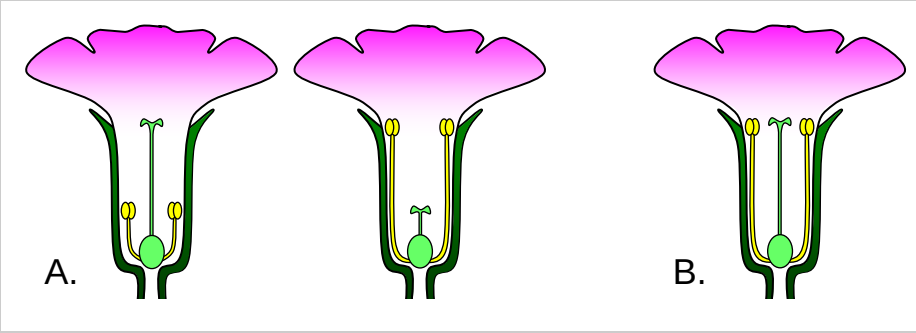
Primula L. (Primulaceae Vent.) consists of ca. 430 species, of which 91% are self-incompatible heterostyles and 9% are self-compatible homostyles (see Box 3; Richards, 2002). Of the homostyles for which the chromosome number is known, 55% are polyploid, whereas only 8% of the heterostylous species are known to be polyploid, suggesting a strong correlation between polyploidy and homostyly (Richards, 2002). In addition, there seems to be a relationship between polyploid homostyles and latitude, especially in *Primula* sect. *Aleuritia* (Richards, 2002).

Box 3. Heterostyly in *Primula

Heterostylous populations are characterised by two floral types (**distyly**) that differ reciprocally from each other in the positions of male (anthers) and female (stigma) sexual organs (see plate A): (i) a long-styled morph with anthers sunken in the corolla tube called '**pin**' (left), and (ii) a short-styled morph with anthers positioned near the mouth of the corolla tube called '**thrum**' (right). In addition to this sex-organ reciprocity called reciprocal herkogamy, heterostylous primulas possess a sporophytically-controlled diallellic self-incompatibility system that prevents selfing and intra-morph mating. The morphological component (i.e. **reciprocal herkogamy**) increases the male fertility by reducing the wastage of pollen through imprecise transfer by pollinators, whereas the physiological component (i.e. **sporophytic self-incompatibility system**) promotes the female fertility by lowering selfing rates, and consequently inbreeding depression.

The heterostylous syndrome is controlled by a cluster of tightly linked, co-adapted genes called the **S supergene**. At least three recombinable loci have been identified so far: (i) the *G* locus controls style length and female incompatibility, (ii) the *P* locus controls pollen size and male incompatibility, and (iii) the *A* locus controls anther height. Thrums are heterozygote (*Ss* or *GPA*/*gpa*) for the genes at the *S* locus and pins are homozygote (*ss* or *gpa*/*gpa*).

Recombination within the S supergene may lead to several variants that rarely occur in the wild, except the **self-fertile long homostyles** (*gPA*) which consist of a single floral morph in which anthers and stigma are of equivalent height and located at the mouth of the corolla tube (see plate B).



*after Lewis and Jones (1992), Richards (1997), and Barrett (2002)

Twenty-seven perennial species are currently ascribed to *Primula* sect. *Aleuritia* Duby (hereafter called *Aleuritia*), 21 to subsect. *Aleuritia* and six to subsect. *Algida* Richards (cf. Fig. 3; Richards, 2002). *Aleuritia* is the only section with species at the *di*ploid (2n=2x=18), *tetra*ploid (2n=4x=36), *hexa*ploid (2n=6x=54), *octo*ploid (2n=8x=72), and *tetradeca*ploid (2n=14x=126) level, and a chromosome base number of x=9 (Richards, 2002). All species known to be diploid are heterostylous and all species known to be polyploid are homostylous, with the exception of the

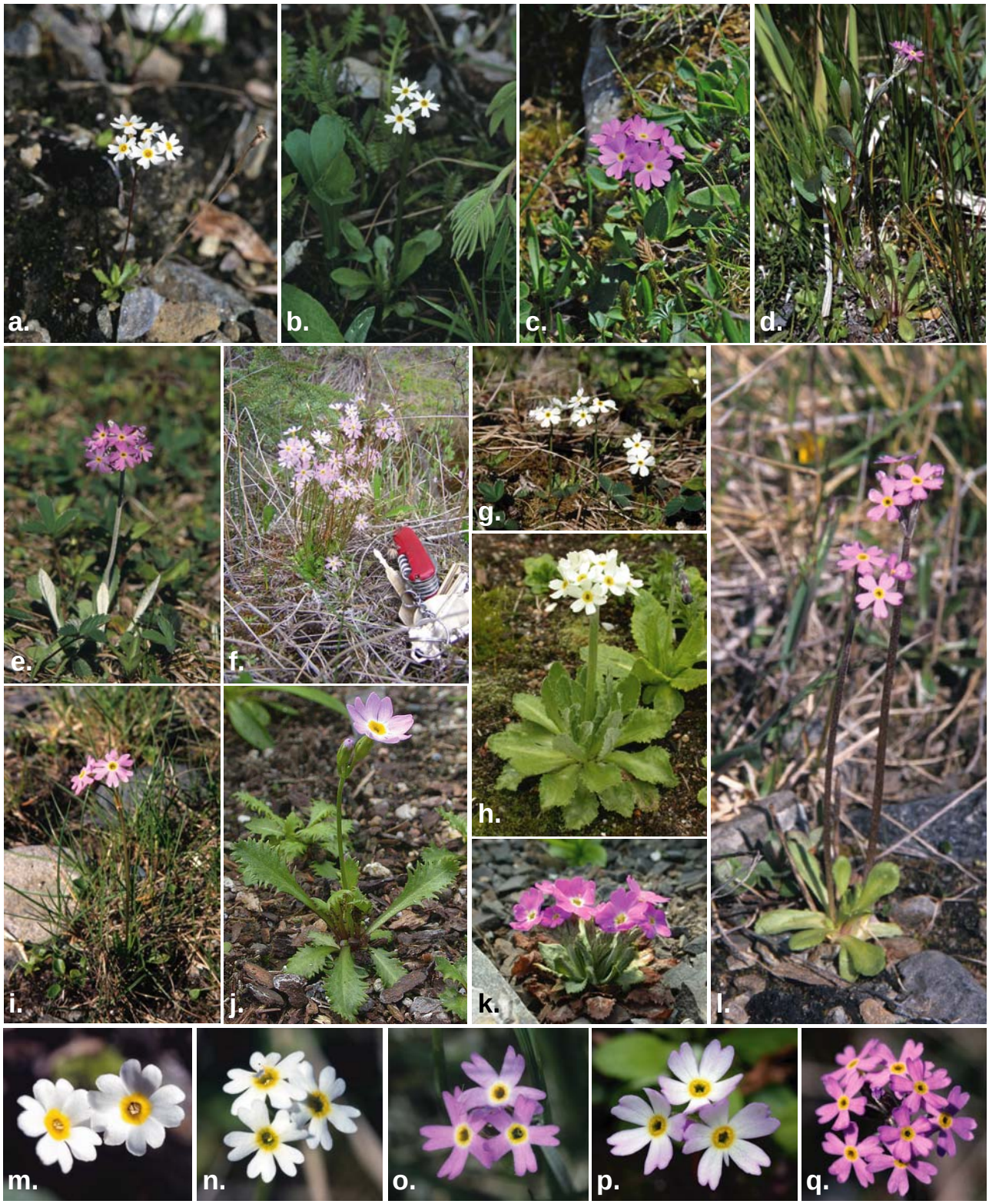


Figure 3. Photographs of selected species of *Primula* sects. *Aleuritia* and the two species of sect. *Armerina* (*P. nutans* and *P. egaliksensis*) supposed to hybridise with *P. mistassinica* of sect. *Aleuritia*. General habit of (a) *P. anvilensis*, (b) *P. egaliksensis*, (c) *P. farinosa*, (d) *P. incana*, (e) *P. laurentiana*, (f-g) *P. mistassinica*, (h) *P. magellanica*, (i) *P. nutans*, (j) *P. pinnata*, (k) *P. yuparensis*, (l) *P. stricta*, and close-up of flowers of (m) *P. anvilensis*, (n) *P. egaliksensis*, (o) *P. incana*, (p) *P. magellanica*, and (q) *P. stricta*. All pictures were taken by A. Guggisberg, except (f) and (j, h, k, p) which were kindly provided by B. Bennett and P. Eveleigh, respectively.

tetraploid *P. borealis* which is heterostyle (Richards, 2002). Members of *Aleuritia* are widely distributed across the boreal and arctic-alpine zone of North-America and Eurasia, and the only species of the genus occurring in the southern hemisphere, *P. magellanica* (Fig. 3h, p), belongs to *Aleuritia* (Richards, 2002). Finally, there is a tendency for polyploids to occur at high latitudes. Indeed, the highest polyploids, i.e. the octoploids *P. laurentiana* (Fig. 3e), *P. scandinavica* and *P. magellanica* (Fig. 3h, p), and the tetradecaploid *P. stricta* (Fig. 3l, q), do not spread south of latitude 45°N or north of latitude 40°S (Richards, 2002).

Aleuritia has been considered as a closely related diploid-polyploid complex, whose origin may be explained within the framework of the secondary contact model proposed by Stebbins (1984; 1985; Fig. 4). According to this model, glacial advancement during the Pleistocene caused the

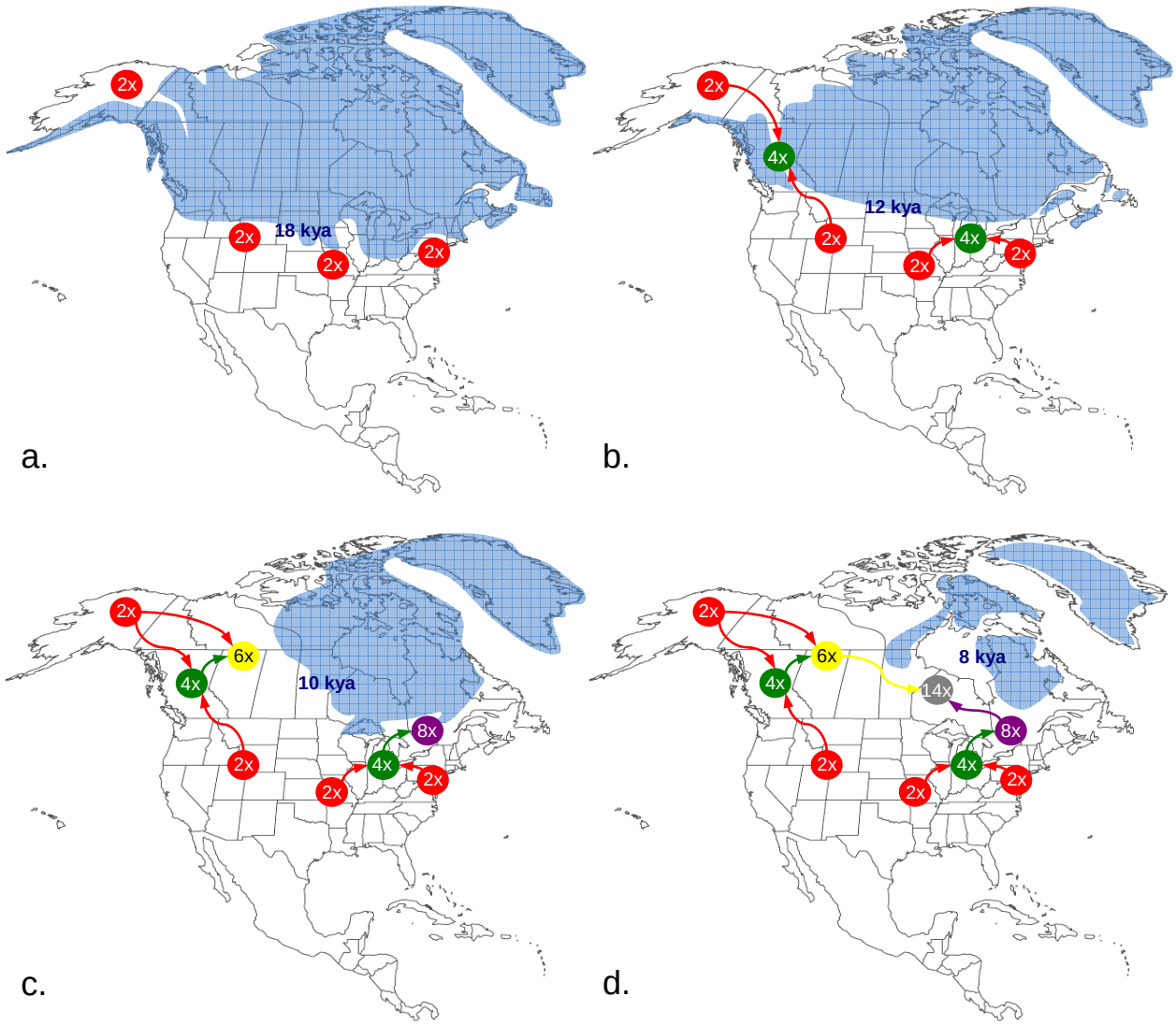


Figure 4. Schematic representation of the secondary contact model. (a) The following map presents the hypothetical distribution of a diploid species (2x; in red) in North-America during the last glacial maximum, ca. 18,000 years ago. The extent of the ice sheets is shown in blue. (b) When the ice sheets start to retreat, some differentiated diploid populations may come into contact again to form tetraploid hybrid complexes (4x; in green). (c) Further retreat of the ice sheets may favour the contact of diploid and tetraploid populations, resulting in the origin of hexaploid hybrids (6x; in yellow), or contact between tetraploid populations may result in the birth of octoploids (8x; in violet). (d) Finally, tetradecaploids (14x; in grey) may result from hybridisation between hexaploid and octoploid populations. kya, thousand years ago.

fragmentation of diploid, allogamous populations, which survived in ice-free areas (i.e. refugia). As glaciers retreated, the differentiated, diploid populations came into contact again and hybridised, giving origin to autogamous, polyploid taxa (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993). The establishment and further success of these newly formed polyploid taxa may be attributed to the switch from heterostyly to homostyly, i.e. the switch from obligate outcrossing to self-compatibility (Baker's law; Baker, 1955; Stebbins, 1957; Hultgård, 1990; Kelso, 1992; Richards, 1997).

In line with the secondary contact model, speciation in *Aleuritia* has been supposed to be driven by hybridisation. Indeed, morphological, caryological and distributional data suggest hybrid origins for most polyploids (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993). It has been further hypothesised that most European polyploids of *Aleuritia* derive from a widespread, diploid *P. farinosa*-like ancestor (Figs. 3c and 5 [in bold]; Bruun, 1932; Hultgård, 1990, 1993), while most American polyploids, including the South-American octoploid *P. magellanica* (Fig. 3h, p), from a widely distributed, diploid *P. mistassinica*-like progenitor (Figs. 3f-g and 5 [in bold]; Kelso, 1991, 1992). Finally, Kelso (1991; 1992) argued for a hybridisation event involving the diploid *P. mistassinica* (Fig. 3f-g) and a diploid species of the Asian sect. *Armerina*, *P. nutans* (Fig. 3i), to explain the origin of the tetraploid *P. egaliksensis* (Fig. 3b, n) of sect. *Armerina* (Fig. 5).

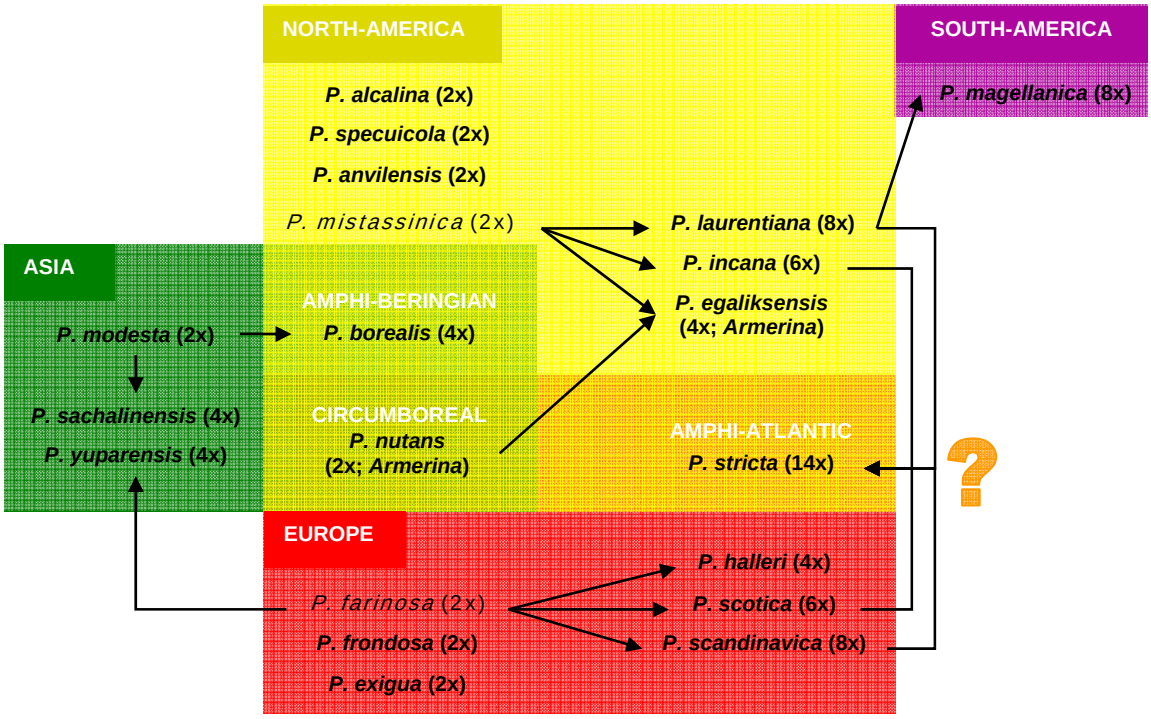


Figure 5. Hypothesised relationships between the 18 species of *Primula* sect. *Aleuritia* subsect. *Aleuritia* of known chromosome number and the two species of sect. *Armerina* (*P. nutans* and *P. egaliksensis*) supposed to hybridise with *P. mistassinica* of sect. *Aleuritia*. Putative ancestor-descendant affinities are indicated by arrows. Source: Bruun (1932), Hultgård (1990; 1993), Kelso (1991), Richards (2002), Vogelmann (1956), www-sbras.nsc.ru/win/elbib/atlas/flora/4129.html

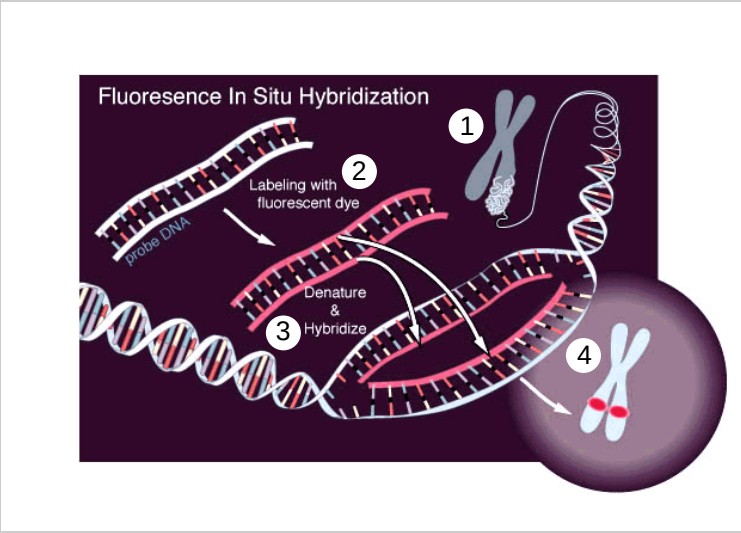
Aleuritia, with its diversity of ploidy levels, breeding systems and distributions represents an ideal model system to study the proposed relationship between polyploidy, self-compatibility and high-latitudinal distribution. Since species of *Aleuritia* concurrently occur in North-America and Eurasia, their study would simultaneously allow testing for parallel patterns in the evolution of these two floras and investigating the potential role of similar climatic and geologic causal factors during the Pleistocene. Clarifying the evolutionary relationships within *Aleuritia* would also contribute to a better understanding of the origin and development of the arctic-alpine flora.

Box 4. Fluorescence and genomic *in situ* hybridisation*

***In situ* hybridisation** (ISH) experiments combine conventional cytogenetics with DNA:DNA hybridisation techniques to physically map specific nucleic acid sequences on morphologically preserved chromosomes. The basic strategy underlying any ISH experiments involves following steps (see plate provided by the National Human Genome Research Institute):

1) Preparation of chromosome spreads: Chromosomes are released from their cells and spread on a glass microscope slide. They constitute the **target DNA**.
2) Labelling of the probe: The DNA fragments designed for hybridisation on the target DNA must be labelled for subsequent localisation (i.e. detection). They constitute the **probe DNA**. When the probe is labelled with a fluorophore, we talk about **fluorescence *in situ* hybridisation** (FISH). **Genomic *in situ* hybridisation** (GISH) is a derivative of FISH, whereby total genomic DNA is used as a probe.
3) DNA:DNA *in situ* hybridisation: To allow annealing between the complementary sequences of the probe and target DNA, both the labelled probe and the chromosomes must be denatured to make them single-stranded.
4) Detection: FISH and GISH signals are localised under epifluorescence light microscopes.

The **stringency** (in percent) at which an ISH experiment is carried out determines the proportion of nucleotides that are correctly matched in the probe and target. For example, at 85% stringency, hybrid molecules with 80% homology will not form. ISH is typically carried out in the range of 70% (described as 'low' stringency) to 90% ('high' stringency) stringency. In practice, the lowest background signal is achieved by hybridising the probe and target at one stringency, and then carrying out post-hybridisation washes at a stringency 2-5% higher.



*after Stace and Bailey (1999), and Schwarzacher and Heslop-Harrison (2000)

Thesis outline

The main goal of this thesis was to disentangle the reticulation patterns in *Primula* sect. *Aleuritia* subsect. *Aleuritia*, and to investigate the consequences of polyploidy on the evolution of breeding systems, using an integrative and multifaceted approach. We began elaborating a maternal phylogeny based on chloroplast sequences to gain a first insight into the biogeographic history of the group, and explore the switches in breeding system, ploidy level and distributional range, to work out an explanation for the success of polyploid taxa in extreme arctic environments (**Chapter 2**). The next step consisted in unravelling the reticulation patterns caused by allopolyploidy and illustrating how the use of different experimental and analytical approaches is fundamental in accomplishing this task. For this purpose, we generated direct and cloned nuclear sequences and gathered evidence from existing morphological, distributional, caryological and phylogenetic data (**Chapter 3**). Lastly, we were interested in verifying the proposed hybrid origin of an intersectional allopolyploid using *in situ* hybridisation techniques (see Box 4), and assess the extent and localisation of genomic reorganisation (recombination) within the parental genomes, to gain preliminary insights into the mechanisms that may be responsible for the switch from heterostyly to homostyly in polyploid taxa (**Chapter 4**).

All chapters summarised above are presented as manuscripts or published papers, and can therefore be read on their own. They are followed by a last chapter (**Chapter 5**) discussing the major findings of the thesis and providing suggestions for future studies.

References

Abbott, R. J., and C. Brochmann. 2003. History and evolution of the arctic flora: in the footsteps of Eric Hultén. Molecular Ecology 12:299-313.

Adams, K. L., and J. F. Wendel. 2005. Novel patterns of gene expression in polyploid plants. Trends in Genetics 21:539-543.

Ainouche, A. K., A. Baumel, and A. Salmon. 2004. *Spartina anglica* C. E. Hubbard: a natural model system for analysing early evolutionary changes that affect allopolyploid genomes. Biological Journal of the Linnean Society 82:475-484.

Arnold, M. L. 1997. Natural Hybridization and Evolution. Oxford University Press, Oxford, UK.

Arnold, M. L. 2006. Evolution Through Genetic Exchange. Oxford University Press, New York, NY, USA.

Baker, H. G. 1955. Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347-349.

Barrett, S. C. H. 1988. The evolution, maintenance, and loss of self-incompatibility systems. Pages 98-124 *in* Plant Reproductive Ecology: Patterns and Strategies (J. Lovett Doust, and L. Lovett Doust, eds.). Oxford University Press, New York, NY, USA.

Barrett, S. C. H. 2002. The evolution of plant sexual diversity. Nature 3:274-284.

Barringer, B. C. 2007. Polyploidy and self-fertilization in flowering plants. American Journal of Botany 94:1527-1533.

Baumel, A., M. Ainouche, R. Kalendar, and A. H. Schulman. 2002. Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C. E. Hubbard (Poaceae). Molecular Biology and Evolution 19:1218-1227.

Bretagnolle, F. 2001. Pollen production and spontaneous polyploidization in diploid populations of *Anthoxanthum alpinum*. Biological Journal of the Linnean Society 72:241-247.

Bretagnolle, F., and J. D. Thompson. 1995. Tansley Review No. 78. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. New Phytologist 129:1-22.

Bretagnolle, F., F. Felber, F. G. Calame, and P. Küpfer. 1998. La polyploïdie chez les plantes. Botanica Helvetica 108:5-37.

Brochmann, C., and R. Elven. 1992. Ecological and genetic consequences of polyploidy in arctic Draba (Brassicaceae). Trends in Ecology and Evolution 6:111-124.

Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. Biological Journal of the Linnean Society 82:521-536.

Bruun, H. G. 1932. Cytological studies in *Primula*, with special reference to the relation between the karyology and taxonomy of the genus. Symbolae Botanicae Upsalienses 1:1-239.

Chen, Z. J., and Z. Ni. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays 28:240-252.

Ehrendorfer, F. 1980. Polyploidy and distribution. Pages 45-60 *in* Polyploidy - Biological Relevance (W. H. Lewis, ed.) Plenum, New York, NY, USA.

Felber, F. 1991. Establishment of a tetraploid cytotype in a diploid population: effect of relative fitness of the cytotype. Journal of Evolutionary Biology 4:195-207.

Fowler, N. L., and D. A. Levin. 1984. Ecological constraints on the establishment of a novel polyploid in competition with its diploid progenitor. The American Naturalist 124:703-711.

Grant, V. 1956. The influence of breeding habit on the outcome of natural hybridization in plants. The American Naturalist 90:319-322.

Grant, V. 1981. Plant Speciation. Columbia University Press, New York, NY, USA.

Harlan, J. R., and J. M. J. DeWet. 1975. On ö winge and a prayer: the origins of polyploidy. Botanical Review 41:361-390.

Henry, Y., M. Bedhomme, and G. Blanc. 2006. History, protohistory and prehistory of the *Arabidopsis thaliana* chromosome complement. Trends in Plant Science 11:267-273.

Hultgård, U.-M. 1990. Polyploidy and differentiation in N European populations of *Primula* subgenus *Aleuritia*. Sommerfeltia 11:117-135.

Hultgård, U.-M. 1993. *Primula scandinavica* and *P. stricta* - patterns of distribution, variation, reproductive strategies and migrations. Opera Botanica 121:35-43.

Husband, B. C., and D. W. Schemske. 1997. The effect of inbreeding in diploid and tetraploid populations of *Epilobium angustifolium* (Onagraceae): implications for the genetic basis of inbreeding depression. Evolution 51:737-746.

Husband, B. C., and H. A. Sabara. 2003. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). New Phytologist 161:703-713.

Husband, B. C., B. Ozimer, S. L. Martin, and L. Pollock. 2008. Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. International Journal of Plant Sciences 169:195-206.

Kelso, S. 1991. Taxonomy of *Primula* sects. *Aleuritia* and *Armerina* in North America. Rhodora 93:67-99.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. Arctic and Alpine Research 24:82-87.

Kovarík, A., R. Matyásek, K. Y. Lim, K. Skalická, B. Koukalová, S. Knapp, M. Chase, and A. R. Leitch. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. Biological Journal of the Linnean Society 82:615-625.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24-40.

Leitch, I. J., and M. D. Bennett. 1997. Polyploidy in angiosperms. Trends in Plant Science 2:470-476.

Leitch, I. J., and M. D. Bennett. 2004. Genome downsizing in polyploid plants. Biological Journal of the Linnean Society 82:651-663.

Levin, D. A. 1975. Minority cytotype exclusion in local plant populations. Taxon 24:35-43.

Levin, D. A. 2002. The Role of Chromosomal Change in Plant Evolution. Oxford University Press, New York, NY, USA.

Levin, D. A., and A. C. Wilson. 1976. Rates of evolution in seed plants: net increase in diversity of chromosome numbers and species numbers through time. Proceedings of the National Academy of Science, USA 73:2086-2090.

Lewis, D., and D. A. Jones. 1992. The genetics of heterostyly. Pages 129-150 *in* Evolution and Function of Heterostyly (S. C. H. Barrett, ed.) Springer Verlag, Berlin, Germany.

Lim, K. Y., R. Matyásek, A. Kovarík, and A. R. Leitch. 2004. Genome evolution in allotetraploid *Nicotiana*. Biological Journal of the Linnean Society 82:599-606.

Liu, B., and J. F. Wendel. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. Molecular Phylogenetics and Evolution 29:365-379.

Liu, B., C. L. Brubaker, G. Mergeai, R. C. Cronn, and J. F. Wendel. 2001. Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome 44:321-330.

Lomolino, M. V., B. R. Riddle, and J. H. Brown. 2006. Biogeography, 3rd edition. Sinauer Associates, Sunderland, MA, USA.

Löve, Á., and D. Löve. 1957. Arctic polyploidy. Proceedings of the Genetics Society of Canada 2:23-27.

Ma, X.-F., and J. P. Gustafson. 2005. Genome evolution of allopolyploids: a process of cytological and genetic diploidization. Cytogenetics and Genome Research 109:236-249.

Madlung, A., and L. Comai. 2004. The effect of stress on genome regulation and structure. Annals of Botany 94:481-495.

Masterson, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264:421-424.

McClintock, B. 1984. The significance of responses of the genome to challenge. Science 226:792-801.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Annual Review of Ecology and Systematics 34:401-437.

Pontes, O., N. Neves, M. Silva, M. S. Lewis, A. Madlung, L. Comai, W. Viegas, and C. S. Pikaard. 2004. Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. Proceedings of the National Academy of Science, USA 101:18240-18245.

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29:467-501.

Rausch, J. H., and M. T. Morgan. 2005. The effect of self-fertilization, inbreeding depression, and population size on autopolyploid establishment. Evolution 59:1867-1875.

Richards, A. J. 1997. Plant Breeding Systems, 2nd edition. Cambridge University Press, Cambridge, UK.

Richards, J. 2002. *Primula*. B. T. Batsford, London, UK.

Rieseberg, L. H., and S. E. Carney. 1998. Plant hybridization. New Phytologist 140:599-624.

Rieseberg, L. H., M. A. Archer, and R. K. Wayne. 1999. Transgressive segregation, adaptation and speciation. Heredity 83:363-372.

Rodríguez, D. J. 1996. A model for the establishment of polyploidy in plants. The American Naturalist 147:33-46.

Schwarzacher, T., and P. Heslop-Harrison. 2000. Practical *in situ* Hybridization. BIOS Scientific Publishers, Oxford, UK.

Segraves, K. A., and J. N. Thompson. 1999. Plant polyploidy and pollination: floral traits and insect visits to diploid and tetraploid *Heuchera grossulariifolia*. Evolution 53:1114-1127.

Skalická, K., K. Y. Lim, M. Matzke, A. R. Leitch, and A. Kovarík. 2005. Preferential elimination of repeated DNA sequences from the paternal, *Nicotiana tomentosiformis* genome donor of a synthetic, allotetraploid tobacco. New Phytologist 166:291-303.

Skalická, K., K. Y. Lim, R. Matyásek, B. Koukalová, A. R. Leitch, and A. Kovarík. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. American Journal of Botany 90:988-996.

Soltis, D. E., and P. S. Soltis. 1993. Molecular data and dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243-273.

Soltis, D. E., and P. S. Soltis. 1999. Polyploidy: recurrent formation and genome evolution. Trends in Ecology and Evolution 14:348-352.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2003. Advances in the study of polyploidy since *Plant Speciation*. New Phytologist 161:173-191.

Soltis, P. S., and D. E. Soltis. 2000. The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Science, USA 97:7051-7057.

Song, K., P. Lu, K. Tang, and T. C. Osborn. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Science, USA 92:7719-7723.

Stace, C. A., and J. P. Bailey. 1999. The value of genomic *in situ* hybridization (GISH) in plant taxonomic and evolutionary studies. Pages 199-210 *in* Molecular Systematics and Plant Evolution (P. M. Hollingsworth, R. M. Bateman, and R. J. Gornall, eds.). Taylor & Francis, London, UK.

Stebbins, G. L. 1950. Variation and Evolution in Plants. Columbia University Press, New York, NY, USA.

Stebbins, G. L. 1957. Self-fertilization and population variability in the higher plants. The American Naturalist 91:337-354.

Stebbins, G. L. 1971. Chromosomal Evolution in Higher Plants. Edward Arnold Ltd., London, UK.

Stebbins, G. L. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. Botanica Helvetica 94:1-13.

Stebbins, G. L. 1985. Polyploidy, hybridization, and the invasion of new habitats. Annals of the Missouri Botanical Garden 72:824-832.

Thompson, J. D., and R. Lumaret. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. Trends in Ecology and Evolution 7:302-307.

Vision, T., D. G. Brown, and S. D. Tanksley. 2000. The origins of genomic duplications in *Arabidopsis*. Science 290:2114-2117.

Vogelmann, H. W. 1956. A Biosystematic Study of *Primula mistassinica* Michx. University of Michigan, Ann Arbor, MI, USA.

Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen, J. J. Lee, B. Watson, T. Kagochi, L. Comai, and Z. J. Chen. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. Genetics 167:1961.

Wendel, J. F. 2000. Genome evolution in polyploids. Plant Molecular Biology 42:225-249.

Chapter 2

Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*

with G. Mansion, S. Kelso, and E. Conti

New Phytologist (2006) 171:617-632

Abstract

Primula sect. *Aleuritia* subsect. *Aleuritia* (*Aleuritia*) includes diploid, self-incompatible heterostyles and polyploid, self-compatible homostyles, the latter generally occurring at higher latitudes than the former. This study develops a phylogenetic hypothesis for *Aleuritia* to elucidate the interactions between Pleistocene glacial cycles, biogeographic patterns, ploidy levels, and breeding systems. Sequences from five chloroplast DNA loci were analysed with parsimony to reconstruct a phylogeny, haplotype network, and ancestral states for ploidy levels and breeding systems. The results supported the monophyly of *Aleuritia* and four major biogeographic lineages: an amphi-Pacific, a South-American, an amphi-Atlantic and a European/North-American lineage. At least four independent switches to homostyly and five to polyploidy were inferred. An Asian ancestor probably gave origin to an amphi-Pacific clade and to a lineage that diversified on the European and American continents. Switches to homostyly occurred exclusively in polyploid lineages, which mainly occupy previously glaciated areas. The higher success of the autogamous polyploid species at recolonising habitats freed by glacial retreat might be explained in terms of selection for reproductive assurance.

Introduction

The evolutionary history of arctic-alpine plants has been deeply influenced by the repeated cooling episodes of the Pleistocene (Dynesius and Jansson, 2000; Abbott and Brochmann, 2003; Hewitt, 2004). A large proportion of arctic-alpine plants are polyploids supposedly derived from hybridisation between diploid progenitors, followed by chromosome doubling – i.e. allopolyploidy (Stebbins, 1950; Löve and Löve, 1975; Brochmann *et al.*, 2004). Many polyploids appear to rely on autogamous fertilisation, while their proposed diploid progenitors are often allogamous, suggesting a causal link between polyploidy and autogamy (Stebbins, 1950; Thompson and Lumaret, 1992). The "secondary contact model", first proposed by Stebbins (1984; 1985) without explicit reference to breeding systems, may be extended to explain the frequent association between polyploidy, autogamy, and current patterns of distribution in arctic-alpine plants. According to this model, glacial advancement during the Pleistocene caused the fragmentation of diploid, allogamous populations, which may have survived in ice-free areas (i.e. refugia). As glaciers retreated, the differentiated diploid populations came into contact again and hybridised, giving origin to polyploid, autogamous taxa. Unreliability of pollinators and/or shifts in pollinator faunas associated with Pleistocene climate cycles might have provided the selective forces that favoured the establishment of the newly-formed polyploid, autogamous species and their higher ability to recolonise deglaciated areas, as compared to their diploid, allogamous progenitors (Stebbins, 1957).

Extensive polyploidisation attributed to cyclic population fragmentation and expansion during the Pleistocene has been documented in several genera, including *Cerastium*, *Draba*, *Parnassia*, *Saxifraga* and *Vaccinium* (reviewed in Abbott and Brochmann, 2003; Brochmann *et al.*, 2004). However, only a few studies have tried to infer the phylogenetic and biogeographic relationships of arctic-alpine polyploid complexes on a global scale. Examples include *Draba* and *Cerastium* (Grundt *et al.*, 2004; Scheen *et al.*, 2004). Furthermore, the combined effects of glacial cycles on ploidy levels and breeding systems have never, to our knowledge, been examined in light of an explicit phylogenetic framework.

Given this background, the study of widespread arctic-alpine polyploid complexes is of special importance in trying to elucidate the interactions between Pleistocene cyclic climate changes, patterns of biotic distribution, polyploidisation, and reproductive biology. *Primula* sect. *Aleuritia* subsect. *Aleuritia* Duby (Primulaceae Vent., – hereafter called *Aleuritia*) represents an ideal case study to investigate such interactions because it is a circum-boreal group that displays variation of ploidy levels, ranging from diploidy to 14-ploidy, and breeding systems, alternating between heterostyly and homostyly. Heterostyly in *Primula* is characterised by two floral morphs (distyly) that differ in the reciprocal positioning of stigma and anthers, usually coupled with a self-

incompatibility system that prevents self and intra-morph fertilisation (Wedderburn and Richards, 1992). Conversely, homostyles are self-compatible (Arnold and Richards, 1998) and may derive from heterostylous progenitors via recombination within the heterostyly linkage group, known as the S supergene (Dowrick, 1956; Wedderburn and Richards, 1992).

Aleuritia, recently supported as monophyletic in genus-wide chloroplast-based phylogenies (Mast *et al.*, 2001; Mast *et al.*, 2006), includes 21 arctic-alpine species of small to medium-sized plants, usually perennial, that typically have a single umbel of flowers, a yellow annulus at the centre of the corolla, presence of dense farina on the stem and calyx, syncolpate pollen, and a base chromosome number of x = 9 (Richards, 2002). Geographically, *Aleuritia* is one of the most widespread groups within *Primula*, with main centres of diversity in the major mountain systems of North-America and Eurasia, except for the Himalayas, and the plains of North-America (Fig. 1a, c-e, Table 1). The only South-American species of *Primula* described so far, *P. magellanica*, belongs to *Aleuritia* (Fig. 1b, Table 1).

Table 1. Distributional range, latitude, ploidy level, breeding system and pollen type of *Primula* sect. *Aleuritia* species.

Primula species	Distributional range	Latitude	Ploidy level	Breeding system	Pollen type
P. modesta Bisset & Moore	Japan, Kurile Is.	33-43°N	2n = 2x = 18	HE	3-4 syncolpate
P. specuicola Rydb.	SE Utah, N Arizona, North Rim, Grand Canyon	36°N	2n = 2x = 18	HE	3-syncolpate
P. exigua Velenovsky	SW Bulgaria	42°N	2n = 2x = 18	HE	na
P. frondosa Janka	Stara Planina (SE Bulgaria)	43°N	2n = 2x = 18	HE	3-syncolpate
P. alcalina Cholewa & Henderson	NE Idaho	45°N	2n = 2x = 18	HE	3-syncolpate
P. mistassinica Michaux	Canada, south to Great Lakes region and west to central Alaska	42-60°N	2n = 2x = 18	HE	3-syncolpate
P. farinosa L.	England, Denmark, Sweden to 64°N, Finnish archipelago; Baltic states to 60°N, Montes Universales (Spain), Pyrenees, Alps, Tatra	43-63°N	2n = 2x = 18	HE	3-syncolpate
P. anvilensis Kelso	Seward Penins., Alaska	65°N	2n = 2x = 18	HE	3-syncolpate
P. incana M. E. Jones	NW America	38-61°N	2n = 6x = 54	HO	4-syncolpate
P. halleri Gmel.	E Alps, Tatra, Carpathians, ex-Yugoslavia, Albania, Rila and Pirin ranges (Bulgaria)	42-47°N	2n = 4x = 36	HO	3-4 syncolpate
P. yuparensis Takeda	Hokkaido (Japan)	43°N	2n = 4x = 36	HO	na
P. magellanica Lehm.	Tierra del Fuego, Patagonia, Falkland Is.	43-55°S	2n = 8x = 72	HO	na
P. laurentiana Fernald	NE America	45-55°N	2n = 8x = 72	HO	4-syncolpate
P. borealis Duby	NW Alaska, N Canada, NE Siberia	52-70°N	2n = 4x = 36	HE	3-syncolpate
P. scotica Hooker	N Scotland	59°N	2n = 6x = 54	HO	4-syncolpate
P. scandinavica (Bruun) Bruun	Norway, a few sites in Sweden	59-70°N	2n = 8x = 72	HO	4-syncolpate
P. stricta Hornem.	Scandinavia, eastward to Kola Penins. (Russia), Novaya Zemlya (Russia), N Iceland, W Greenland, NE Canada	62-73 °N	2n = 10-16x = ca. 88-136	HO	5-syncolpate
P. baldschuanica B. Fedstch.	Tadzhikistan, E Afghanistan	na	na	na	na
P. capitellata Boiss.	Iran, Afghanistan, Pakistan	na	na	HE	3-syncolpate
P. sachalinensis Nakai	E Siberia, Kamtschatka and Sakhalin I.	na	na	na	na
P. schlagintweitiana Pax	N Pakistan, Kashmir and NW India	na	na	na	na

Species included in the present study are highlighted in boldface. HE, heterostyle; HO, homostyle; na, not available.
Source: (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991; Hultgård, 1993; Richards, 2002)

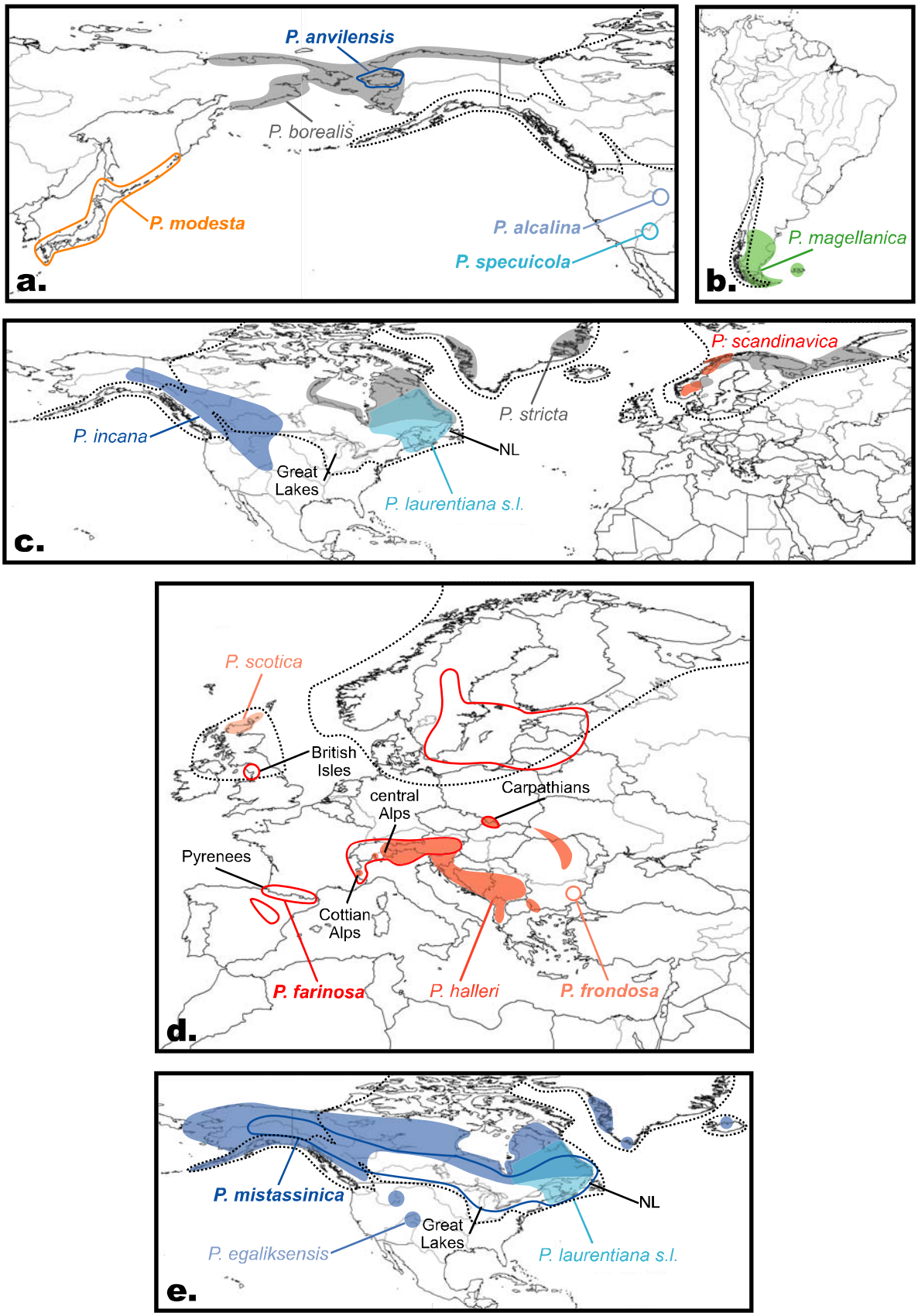


Figure 1. Distributional ranges of taxa from *Primula* sects. *Aleuritia* and *Armerina* discussed in this study. Ranges of diploids (in bold type) are shown by plain contour lines, those of polyploids by shaded areas. Ranges of Asian, South-American, North-American and European species are tinged in orange, green, blue and red tones, respectively, whereas species occurring on more than one continent are shown in grey tones. Dashed, black lines symbolise the extent of major ice sheets during the last glacial maximum (18 000BP, according to Siegert, 2001). Species ranges are grouped according to the clades of Figure 2: (a) clade A; (b) clade B; (c) clade C; (d) sub-clade d_1; (e) sub-clade d_2; NL, Newfoundland.

The *Aleuritia* polyploid system comprises at least five different ploidy levels, including the diploids *P. alcalina*, *P. anvilensis*, *P. exigua*, *P. farinosa*, *P. frondosa*, *P. mistassinica*, *P. modesta* and *P. specuicola* (2n = 18), the tetraploids *P. borealis*, *P. halleri* and *P. yuparensis* (2n = 36), the hexaploids *P. incana* and *P. scotica* (2n = 54), the octoploids *P. laurentiana*, *P. magellanica* and *P. scandinavica* (2n = 72), and the 14-ploid *P. stricta* (2n = ca. 88-136; Table 1). Chromosome numbers are unavailable for four Asian endemics: *P. baldschuanica, P. capitellata, P. sachalinensis* and *P. schlagintweitiana*. All eight diploid species are heterostylous, whereas eight of the nine polyploid species are homostylous. The tetraploid *P. borealis* is heterostylous. Furthermore, the highest polyploids (8x-14x) rarely occur south of latitude 45°N or north of latitude 40°S, whereas the diploids usually occur between these latitudes (Fig. 1, Table 1). This geographic pattern suggested an association between high-latitudinal distribution, polyploidy, and homostyly (Bruun, 1932; Richards, 2002) and has been explained within the framework of the secondary contact model described above (Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993).

Despite the publication of several recent papers on *Primula* (e.g. Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Conti *et al.*, 2000; Mast *et al.*, 2001; Trift *et al.*, 2002; Mast *et al.*, 2006), no study has attempted to elucidate the evolutionary history and biogeography of *Aleuritia* within a detailed phylogenetic framework. Furthermore, while the proposed correlation between homostyly and polyploidy has been investigated in other genera, e.g. *Amsinckia*, *Damnacanthus* and *Turnera* (Barrett and Shore, 1987; Schoen *et al.*, 1997; Naiki and Nagamasu, 2004; Truyens *et al.*, 2005), the combined effects of glacial cycles on these traits and biotic patterns of distribution has never been explored in light of an explicit phylogenetic hypothesis. Therefore, the present study intends to reconstruct the phylogenetic relationships within *Aleuritia* to (i) elucidate the biogeographic history of the European and American taxa, (ii) infer whether ploidy levels and breeding systems changed once or rather multiple times, and (iii) interpret the results in light of the selective forces that are likely to have modulated the interconnected evolution of ploidy levels, reproductive biology, and biotic ranges.

Materials and Methods

Taxon and locus sampling

A total of 63 plant accessions were used for the present study. Forty-four accessions represented 15 of the 21 species ascribed to *Aleuritia* and 19 accessions represented four of the 14 species ascribed to sect. *Armerina* (Richards, 2002). The *Armerina* accessions were included because cytological, morphological and biogeographic evidence suggested that *P. egaliksensis*, assigned to sect. *Armerina* (Richards, 2002), might have a hybrid origin involving an *Aleuritia* and an *Armerina* parent (Kelso, 1991). Furthermore, a phylogeny of *Primula* based on two chloroplast introns

supported the inclusion of selected members of *Armerina* and *Aleuritia* in the same, largely unresolved clade and, more specifically, of *P. egaliksensis* within a well-supported *Aleuritia* sub-clade (Mast *et al.*, 2001; Mast *et al.*, 2006). Finally, *P. ioessa* (sect. *Sikkimensis*) was chosen to root the resulting tree because it was a member of the same large clade comprising sects. *Aleuritia* and *Armerina* in the mentioned *Primula* phylogeny (Mast *et al.*, 2001; Mast *et al.*, 2006).

Sequences were generated from five non-coding regions of the chloroplast DNA (cpDNA), namely the *rpl*16 intron, *rps*16 intron, *trn*L intron, *trn*L-F spacer, and *trn*T-L spacer (see Appendix 1). The maternal inheritance of the cpDNA has been demonstrated in many angiosperms, including *Primula* (Corriveau and Coleman, 1988).

DNA extraction, PCR amplification, sequencing

Total genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Switzerland), after grinding the leaf material with glass beads in a Retsch MM 2000 shaker (Retsch, Germany).

Polymerase chain reactions (PCR) were performed in 20 µl volumes containing 1X buffer (including 1.5 mM $MgCl_2$), 2 mM $MgCl_2$, 200 µM dNTPs, 0.2 µM of each primer, and one unit Taq polymerase (Sigma, Switzerland). Amplifications were carried out on a thermocycler (Biometra, Germany), using the following conditions: a first cycle at 94°C for 2 min; 35 cycles at 94°C for 30 s, 52°C for 1 min and 72°C for 1.75 min; a final cycle of 10 min at 72°C. The *rpl*16 intron was amplified with primers rpL16F71 (5'-GCTATGCTTAGTGTGTGACTCGTTG-3') and rpL16R1516 (5'-CCCTTCATTTCTTCCTCTATGTTG-3'; Small *et al.*, 1998), the *rps*16 intron with primers rpS16F (5'-GTGGTAGAAAGCAACGTGCGACTT-3') and rpS16R2 (5'-TCGGGATCGAACATCAATTGCAAC-3'; Oxelman *et al.*, 1997), the *trn*L intron and *trn*L-F spacer – hereafter called *trn*L-F region – with primers trnL5'^{UAA}F (TabC, 5'-CGAAATCGGTAGACGCTACG-3') and trnF^{GAA} (TabF, 5'-ATTTGAACTGGTGACACGAG-3'; Taberlet *et al.*, 1991), and the *trn*T-L spacer with primers trnT^{UGU}F (TabA, 5'-CATTACAAATGCGATGCTCT-3') and 5'trnL^{UAA}R (TabB, 5'-TCTACCGATTTCGCCATATC-3'; Taberlet *et al.*, 1991). The length of the amplified fragments was estimated by comparison with DNA ladders on 1.2% agarose gels stained with ethidium bromide. The absence of contamination was checked by running out the PCR products of a negative control on the same gel. Successfully amplified PCR products (amplicons) were cleaned with the GFX PCR DNA and Gel Band Purification Kit (Biosciences Amersham, Switzerland).

Sequencing reactions were prepared with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA), using the same primers as in the PCR amplifications. Sequencing products were purified on 96-well multiscreen filtration plates (Millipore, USA) to remove excess dye terminators, and run on an ABI Prism 3100 automated

sequencer (Applied Biosystems, USA). Sequencher 4.2 (Gene Codes Corp., USA) was used to check the quality of the electropherograms and compile the contiguous sequences (contigs) for each amplicon.

Phylogenetic analyses

The starting and ending points of each sequence were determined by comparison with the complete cpDNA sequence of *Nicotiana tabacum* (GenBank Z00044). Nucleotide sequences were aligned by eye (see Table 2, columns a). Unequivocally aligned gaps were coded as present or absent with the software GapCoder (Young and Healy, 2003), and added as binary characters to the end of the four-state nucleotide matrix (see Table 2, columns b).

Each data set was first analysed separately under maximum parsimony (MP) optimisation. Visual comparisons of the trees derived from each data set independently revealed no strongly supported (bootstrap values > 59%) topological incongruence among the different cladograms. Therefore, data partitions were combined into a "global matrix" that was analysed with MP using the beta 10 version of PAUP* 4.0 (Swofford, 1999). Heuristic searches were performed with character states weighted equally, gaps treated as missing data, and the following parameters: TBR branch swapping, Steepest Descent ON, Mulpars ON, and Collapse branches option ON for branches with a minimum length of zero. Two hundred searches were performed under these conditions, after randomising the order of taxon addition. Five trees per replicate were saved and used as starting trees for a further round of branch swapping with TBR, now saving all the trees. Branch support evaluation was performed using the bootstrap method (Felsenstein, 1985). Bootstrap values (BS) were obtained from 100,000 fast-bootstrap replicates. This method is known to produce BS that are generally lower than those obtained using heuristic searches with stepwise addition and branch swapping (DeBry and Olmstead, 2000; Mort *et al.*, 2000).

Haplotype network

Conventional methods of phylogenetic reconstruction often produce low levels of resolution in cases of insufficient variability (Crandall, 1994), reticulation (e.g. hybridisation), or multifurcation (i.e. single ancestral haplotypes giving rise to multiple descendants, Posada and Crandall, 2001). Under such circumstances, network analyses such as statistical parsimony, implemented in TCS 1.21 (Clement *et al.*, 2000), can be profitably applied (e.g. Verheyen *et al.*, 2003; Lihová *et al.*, 2004). This analytical approach produces unrooted networks connecting only the haplotypes that have a high probability (> 0.95) of being similar due to shared history and not homoplasy (i.e. multiple hits, Templeton *et al.*, 1992). Analyses were run on the global matrix, including all

generated sequences, but excluding indels and missing data, which may create errors when collapsing sequences into haplotypes.

Character-state reconstructions

Patterns of character evolution were reconstructed using version 4.0 of MacClade (Maddison and Maddison, 2000). To perform these analyses, the strict consensus tree of Figure 2 was first pruned according to the following criteria: (1) when multiple accessions of the same species were monophyletic, only a single, randomly selected sequence was kept; (2) when multiple accessions of the same species were non-monophyletic (see *P. laurentiana* and *P. stricta*), a single randomly selected sequence of each clade was kept; (3) *P. halleri* and *P. scotica*, which formed a polytomy and shared the same character states, were merged into a single taxonomic unit (halleri/scotica 1-2; see Fig. 4); (4) *P. egaliksensis* and two accessions of *P. laurentiana*, which formed a polytomy and shared the same character states, were merged into a single taxonomic unit (egaliksensis 1-8/laurentiana 3-4; see Fig. 4). Consequently, the tree used for character state reconstructions included 20 terminals. Ancestral state optimisations were performed on two topologies that differed in the position of *P. magellanica*, to reflect the different relationships of this taxon in the MP trees derived from the global matrix (data not shown).

The following two binary characters were evaluated: (1) breeding system (a. heterostyly, b. homostyly), and (2) ploidy level (a. diploidy, b. polyploidy). For character 1, the unordered states 1a and 1b were either equally weighted (i.e. a change from heterostyly to homostyly or vice-versa had the same cost of one) or differentially weighted with a 2:1 ratio, favouring the loss of heterostyly over its gain. This weighting strategy was adopted because it has been argued that the origin of a complex trait, such as heterostyly, is likely to be less common than its loss (Kohn *et al.*, 1996; Schoen *et al.*, 1997). Likewise, we investigated the effects of using equal weights vs. a 2:1 weighting ratio (favouring the shift from diploidy to polyploidy) on character state optimisations for character 2. The weighting ratio was selected because polyploids may become established (i.e. fully fertile) after only a few generations (Ramsey and Schemske, 2002), while supposedly requiring a much longer time span to get fully "diploidised" again (Wendel, 2000; Leitch and Bennett, 2004). Both delayed (DELTRAN) and accelerated (ACCTRAN) transformations were used to resolve equivocal reconstructions.

To explicitly test the proposed correlation between polyploidy and homostyly within *Aleuritia*, we used the concentrated-changes test (CCT) implemented in MacClade 4.0 (Maddison, 1990). This test evaluates whether changes in two binary characters are randomly dispersed on a tree or rather concentrated along specific branches. More specifically, it tests whether gains or losses in a dependent character are more concentrated on the branches of an independent character than

expected by chance (Maddison and Maddison, 2000). The probability of concentrated changes was assessed by exact counts in the *Aleuritia* clade for each reconstruction listed in Table 3. The same optimisation scheme was used for both characters.

We designated character 2 (ploidy level) as the independent variable and character 1 (breeding system) as the dependent variable, based on the following rationale. In some plant taxa, polyploid genomes have been demonstrated to undergo rapid genomic rearrangements after their formation (Wendel, 2000; Soltis *et al.*, 2003). This higher rate of recombination might increase the probability that the heterostyly linkage group is disrupted, thus explaining the higher frequency of homostyly in polyploids, as suggested by previous studies on *Primula* (Dowrick, 1956; Wedderburn and Richards, 1992). In this context, polyploidisation would favour homostyly, justifying the selection of ploidy level as the independent variable. Alternatively, homostyly might have preceded polyploidisation, but this hypothesis is contradicted by the absence of diploid homostyles and the occurrence of a polyploid heterostyle, *P. borealis*, in *Aleuritia* (Table 1).

Results

Phylogenetic analyses

The main characteristics of the five cpDNA regions used for the present study, along with the corresponding trees statistics, are summarised in Table 2. Parsimony analyses performed on the global data set, including gaps in the aligned matrix, produced four MP trees (length [L] = 335, consistency index [CI] = 0.928, retention index [RI] = 0.981, rescaled consistency index [RC] = 0.911; Fig. 2). The addition of coded gaps to the global data set resolved five additional clades and increased BS values for four branches, while causing a decrease only for the clade comprising all accessions of *P. mistassinica* (Table 2, Fig. 2). Some *P. mistassinica* accessions differ from each other by only one change, which is shared with one accession of *P. laurentiana* in a different clade (laurentiana 4).

The strict consensus of the four MP trees generated from the global data set supported the monophyly of sections *Aleuritia* and *Armerina* (BS 100%; Fig. 2), except for the placement of *P. egaliksensis* (taxonomically assigned to sect. *Armerina*) in the *Aleuritia* clade. *Aleuritia* can be subdivided into four main clades. Clade A (BS 92%), sister to the rest of *Aleuritia*, comprises four diploid species, *P. modesta*, *P. specuicola*, *P. alcalina* and *P. anvilensis*, and one tetraploid species, *P. borealis*. Clades B–D form a well-supported super-clade (BS 100%), but remain unresolved with respect to each other (Fig. 2). Clade B (BS 97%) comprises three accessions of the South-American octoploid, *P. magellanica* (BS 97%). The poorly supported clade C (BS 63%) consists of several European and North-American polyploids, namely *P. stricta* (14x), *P. scandinavica* (8x), *P. incana* (6x), and two accessions of *P. laurentiana* (8x; laurentiana 1-2). Finally, the main clade D (BS

98%) consists of a European and a North-American sub-clade (sub-clades d_1 and d_2, respectively). The European sub-clade d_1 is weakly supported (BS 59%), and the relationships among its taxa remain largely unresolved, as very low levels of cpDNA variation were detected between *P. farinosa* (2x), *P. frondosa* (2x), *P. halleri* (4x) and *P. scotica* (6x). Conversely, the North-American sub-clade d_2, formed by *P. mistassinica* (2x), *P. egaliksensis* (4x; sect. *Armerina*), and two accessions of *P. laurentiana* (8x; laurentiana 3-4), is strongly supported (BS 86%).

Table 2. Summary statistics of the five chloroplast loci used for this study.

	*rpl*16 intron		*rps*16 intron		*trn*L-F region	
	a	b	a	b	a	b
Character metrics:						
total (gaps)	1003	1026 (23)	860	873 (13)	1058	1088 (30)
informative (gaps)	27	36 (9)	29	34 (5)	23	36 (13)
uninformative (gaps)	26	40 (14)	24	32 (8)	26	43 (17)
PICs (gaps)	53	76 (23)	53	66 (13)	49	79 (30)
% variability	5.28	7.39	6.16	7.56	4.63	7.26
Tree statistics:						
N	1	2	1	1	1	4
L	59	85	58	72	52	89
CI	0.9492	0.929	0.9091	0.9444	0.96	0.91
RI	0.9881	0.979	0.9878	0.9859	0.9957	0.976
RC	0.9379	0.91	0.9367	0.9311	0.9765	0.888
Effects of including gaps:						
additional clades		3		none		3
no. of branches with higher BS		none		2		1

Character metrics (a, matrix with nucleotides only; b, matrix with nucleotides and coded gaps): total length of matrix (number of gaps); number of informative characters in the aligned matrix (number of informative gaps); number of uninformative characters in the aligned matrix (number of uninformative gaps); number of potentially informative characters or PICs (number of potentially informative gaps); and percent variability, calculated by dividing the PICs value by the length of the aligned matrix. Tree statistics: number of trees (N); tree length (L); consistency index (CI); retention index (RI); and rescaled coefficiency index (RC). The effects of including gaps as additional characters on phylogenetic resolution are summarised at the end of the table.

Table 2 (continued)

	*trn*T-L spacer		Global dataset	
	a	b	a	b
Character metrics:				
total (gaps)	884	900 (16)	3805	3887 (82)
informative (gaps)	34	41 (7)	113	146 (34)
uninformative (gaps)	31	40 (9)	107	156 (48)
PICs (gaps)	65	81 (16)	220	302 (82)
% variability	7.35	9	5.78	7.77
Tree statistics:				
N	1	1	4	4
L	71	87	242	335
CI	0.944	0.954	0.9015	0.928
RI	0.987	0.989	0.9876	0.981
RC	0.932	0.943	0.9345	0.911
Effects of including gaps:				
additional clades		3		5
no. of branches with higher BS		1		4

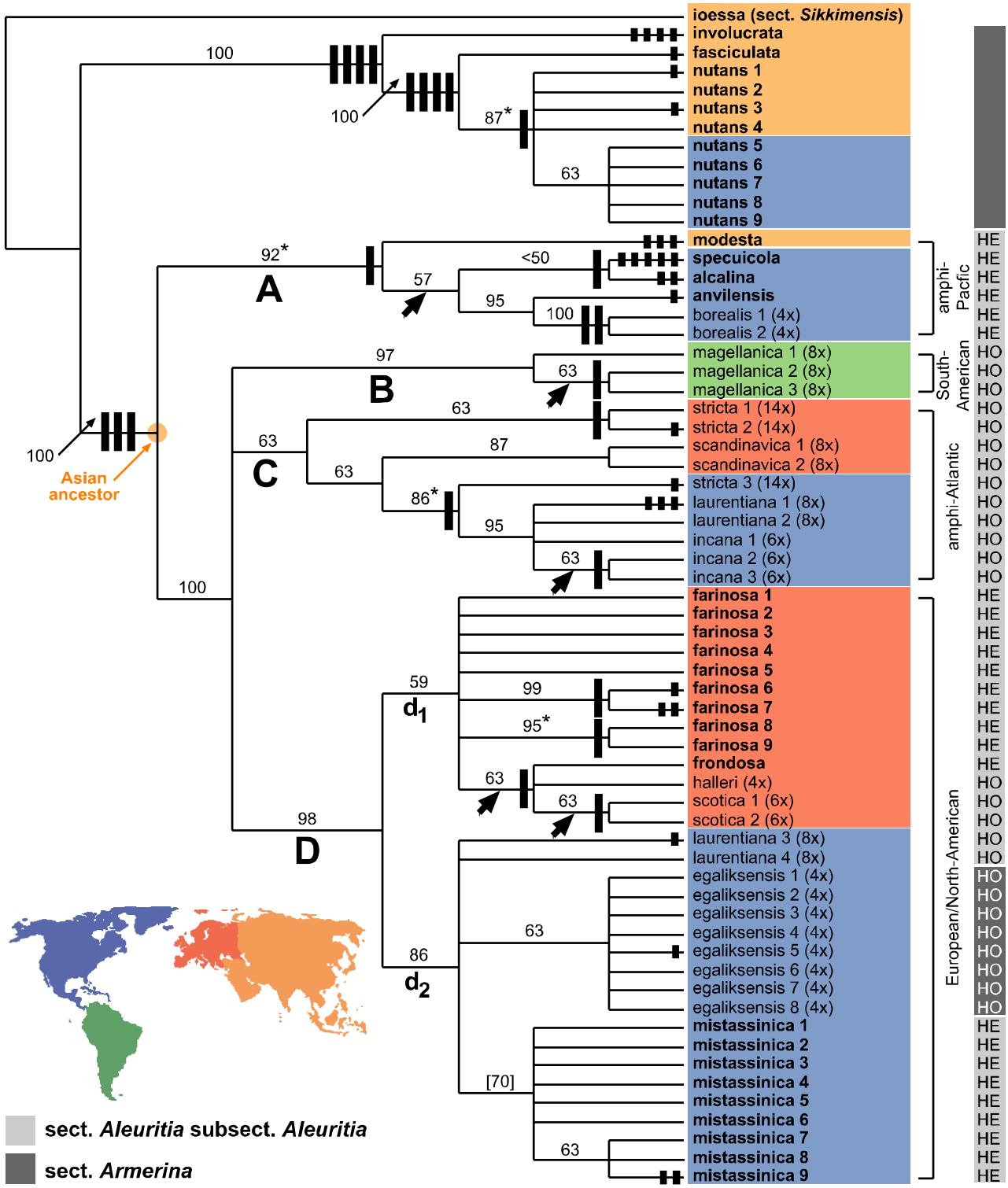


Figure 2. Strict consensus of the four most parsimonious trees from the analysis of the global matrix. Diploid taxa are in bold type; polyploids are followed by the ploidy levels reported in the literature (see Table 1). Asian, South-American, North-American and European accessions are coloured in orange, green, blue and red, respectively. Bootstrap support (BS) values over 50% are shown above the branches. Asterisks identify branches with higher BS values after adding coded gaps to the data matrix, brackets those with lower BS values, and thick arrows additional clades. Non-homoplasious gaps are mapped; synapomorphic ones are represented by long bars, autapomorphic ones by short bars. Gaps supporting the ingroup are not mapped. Major clades are identified by capital letters (A-D). Abbreviations for terminal taxa are given in Appendix 1. HE, heterostyle; HO, homostyle.

Haplotype network

Seven species, for which multiple accessions were available, showed some degree of infraspecific variation. Two haplotypes were detected within *P. scotica* and *P. egaliksensis*, respectively, while three species, *P. laurentiana*, *P. mistassinica* and *P. stricta*, included three haplotypes each. Finally, five haplotypes were found within both *P. farinosa* and *P. nutans* (sect. *Armerina*). While most infra-specific haplotypes differed only by a few nucleotides, the four haplotypes of *P. laurentiana* differed so much from each other that they were placed in unrelated clades (C and d_2, respectively; Figs. 2-3).

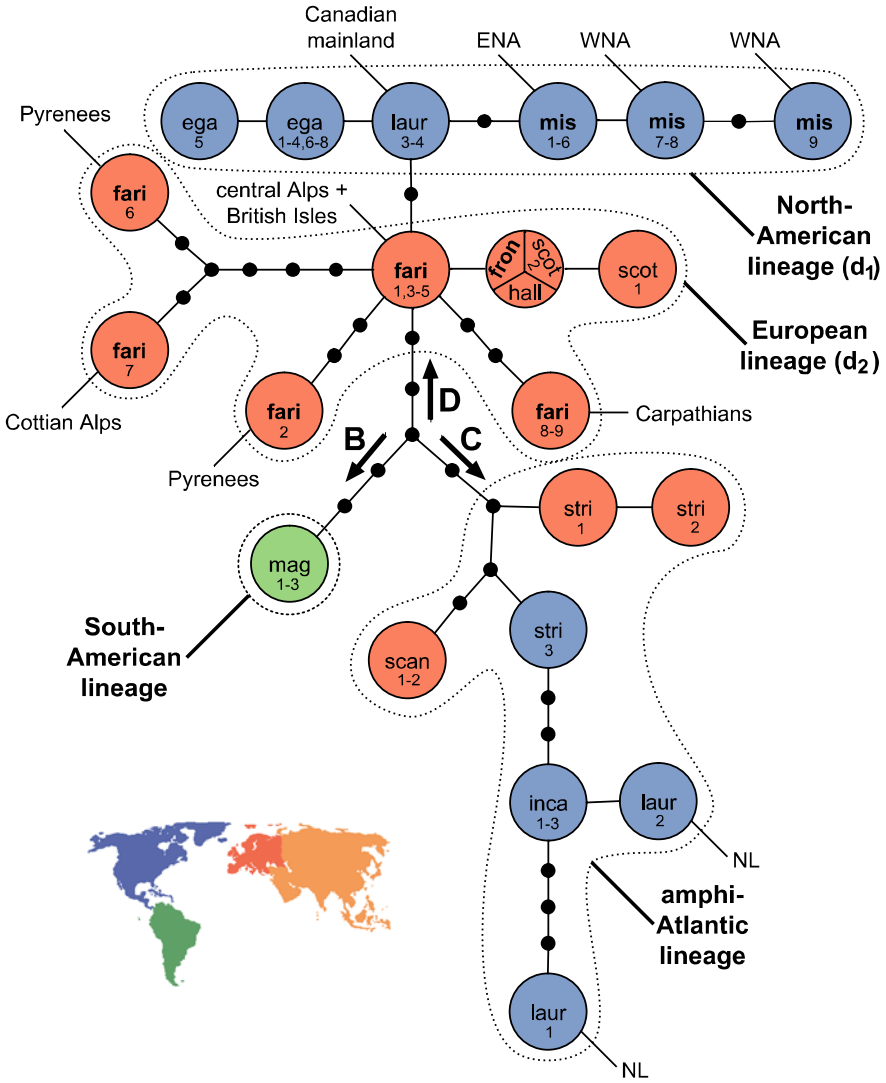


Figure 3. Haplotype network for clades B, C, and D of Figure 2 derived from statistical parsimony. Arrows indicate divergence from the hypothetical most ancestral haplotype. Diploid taxa are in bold. Asian, South-American, North-American and European accessions are coloured in orange, green, blue and red, respectively. Lines within the network represent single mutational steps, and small circles missing haplotypes. Accessions numbered as in Figure 2 (see Appendix 1 for further details). A dash between accession numbers means "from haplotype number … to haplotype number …", a comma means "haplotype number … and haplotype number …"; e.g. "farinosa 1, 3-5" means "haplotype 1, and haplotypes 3 to 5 of *P. farinosa*". ega, *P. egaliksensis*; fari, *P. farinosa;* fron, *P. frondosa;* hall, *P. halleri*; inca, *P. incana*; laur, *P. laurentiana*; mag, *P. magellanica*; mis, *P. mistassinica*; scan, *P. scandinavica*; scot, *P. scotica*; stri, *P. stricta*. Geographic areas discussed in the text are labelled congruently with the labels used in Figure 1. ENA, eastern North-America (i.e. east of Great Lakes); NL, Newfoundland; WNA, western North-America (i.e. west of Great Lakes).

The statistical parsimony analysis performed on the global matrix collapsed the sequences to 34 haplotypes and recovered two unlinked networks. The first network connected 21 haplotypes representing 45 accessions (Fig. 3). The haplotypes were linked into three major groups, corresponding to clades B, C and D of the MP consensus tree (see Fig. 2). Three mutational steps linked each one of these groups to the centre of the network, which supposedly represents a missing ancestral haplotype connecting groups B, C, and D to clade A (see Fig. 2). Group B connects the three identical accessions of *P. magellanica*; group C connects all accessions of *P. incana*, *P. scandinavica* and *P. stricta* along with two samples of *P. laurentiana* (laurentiana 1-2); group D connects all specimens of *P. farinosa*, *P. frondosa*, *P. halleri*, *P. scotica*, *P. egaliksensis*, and *P. mistassinica*, along with two identical accessions of *P. laurentiana* (laurentiana 3-4). Samples of *P. farinosa* collected in the central Alps (farinosa 1, 3) and England (farinosa 4-5), form the central haplotype of group D, called the "alpine" haplotype from now on (Fig. 3).

Character-state reconstructions

The different position of *P. magellanica* in the two MP trees obtained from the reduced data set used for character state reconstructions had no effect on ancestral state optimisations for either character 1 (breeding system) or 2 (ploidy level). All reconstructions supported heterostyly as the ancestral state in *Aleuritia*, regardless of the used weighting scheme or optimisation strategy. However, the number of times that homostyly was gained and lost changed with different weighting schemes and optimisation strategies (Table 3). For example, under equal weighting (1:1 ratio), three gains and one loss of homostyly were reconstructed with ACCTRAN optimisation, while four gains of homostyly were inferred when DELTRAN was used to resolve ambiguous optimisations. When shifts from heterostyly to homostyly were considered twice as probable as the alternative change, character state optimisation for breeding systems was identical to that inferred under equal weights and DELTRAN (Fig. 4).

Similarly to heterostyly, all reconstructions supported diploidy as the ancestral state in *Aleuritia*, regardless of the used weighting scheme or optimisation strategy. However, the number of switches from diploidy to polyploidy (i.e. gains) varied according to the applied weighting scheme and strategy used to resolve ambiguous optimisations (Table 3). Under equal weighting (1:1 ratio), four switches to polyploidy and one reversal to diploidy were reconstructed with ACCTRAN, whereas five polyploidisation events were inferred with DELTRAN. Finally, when a 2:1 weight was employed, the pattern of character state reconstructions for ploidy level was identical to the one obtained with equal weights and DELTRAN (Fig. 4).

In the reconstructions shown in Figure 4, both obtained with a 2:1 weighting scheme, four origins of homostyly were inferred along the five branches of the polyploid lineages, rejecting the

null hypothesis that changes are distributed randomly ($P < 0.05$). Alternative reconstructions, which were inferred using an equal weighting scheme and ACCTRAN/DELTRAN optimisations (not shown), yielded similar results.

Table 3. Summary of character state reconstructions

	Weighting scheme	Resolving options	Gains (a → b)	Losses (b → a)
Character 1 (breeding system):	1:1 ratio	ACCTRAN	3	1
a. heterostyly, b. homostyly	1:1 ratio	DELTRAN	4	0
	2:1 ratio	no ambiguous reconstructions	4	0
Character 2 (ploidy level):	1:1 ratio	ACCTRAN	4	1
a. diploidy, b. polyploidy	1:1 ratio	DELTRAN	5	0
	2:1 ratio	no ambiguous reconstructions	5	0

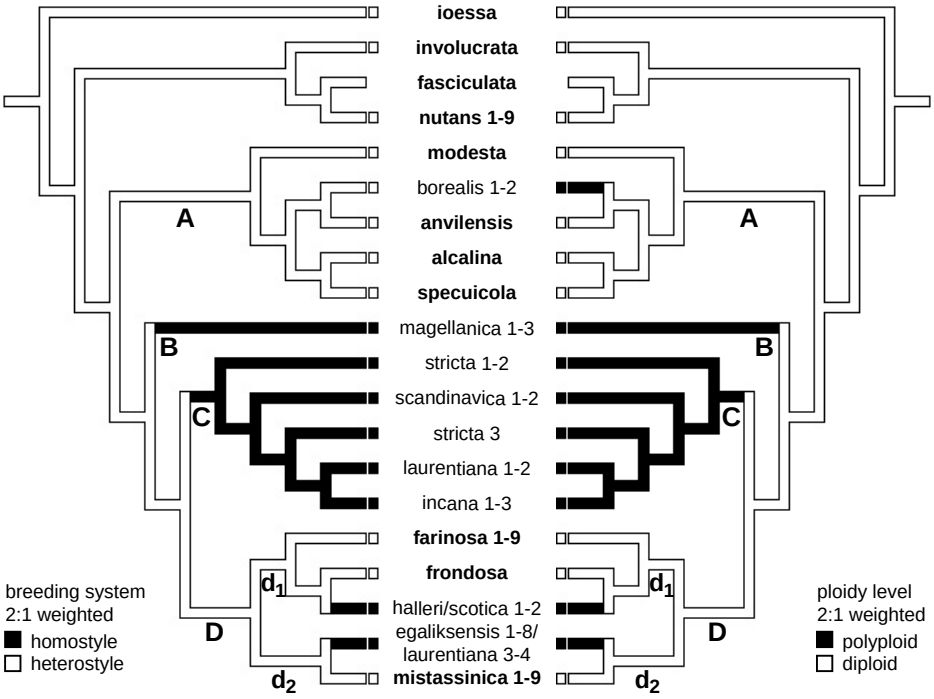


Figure 4. Ancestral states reconstructions for breeding system (left) and ploidy level (right) under 2:1 weighting schemes, favouring the loss of distyly and diploidy, respectively (see Table 3). Diploid taxa are in bold and major clades as in Figure 2. Abbreviations for terminal taxa are given in Appendix 1.

Discussion

The phylogenetic results presented here provide novel evidence to explore the effects of Pleistocene glacial cycles on the evolution of biogeographic patterns, ploidy levels, and breeding systems in the *Aleuritia* diploid-polyploid species complex. The phylogeny, based on the maternally inherited chloroplast genome, also allows us to suggest the possible maternal lineages that contributed to the origins of polyploids in *Aleuritia*.

Biogeographic patterns and ploidy levels in Aleuritia

The cpDNA phylogeny confirms the monophyly of *Aleuritia*, including *P. egaliksensis* (taxonomically ascribed to sect. *Armerina*, Richards, 2002), and supports four main clades, discussed below (Figs. 2-3). The sister group relationship of the mainly Asian section *Armerina* to

Aleuritia (Fig. 2), the topology of a genus-wide cpDNA phylogeny of *Primula* (Mast *et al.*, 2001; Mast *et al.*, 2006), and the observation that the Himalayan mountain system represents the main centre of species diversity for *Primula* (Richards, 2002), suggest that an Asian ancestor possibly gave origin to an amphi-Pacific clade and to a lineage that diversified on the European and American continents. This lineage may have spread westward across the central Asian mountain ranges or via the arctic shorelines, giving origin to the amphi-Atlantic and European/North-American clades, and south-eastward, giving origin to the sole South-American species of *Primula* (*P. magellanica*). Clearly, a more rigorous interpretation of the geographic origin of *Aleuritia* will require an explicit reconstruction of ancestral areas of distribution at the genus-wide level. Furthermore, the haplotype network (Fig. 3) allowed us to identify the central biogeographic role of the "alpine" haplotype of *P. farinosa* in Europe and suggest that it might have given origin to the North-American accessions of clade D.

Amphi-Pacific lineage (clade A)

This strongly supported clade (BS 92%, Fig. 2), sister to the rest of *Aleuritia*, consists of four diploids (*P. modesta*, *P. anvilensis*, *P. specuicola*, and *P. alcalina*) with very narrow ranges and one widespread tetraploid (*P. borealis*). The phenotypically variable *P. modesta* ranges from Japan to the Kurile Islands, with a few disjunct localities on the Korean mainland and neighbouring Siberia (Fig. 1a). *Primula anvilensis* is restricted to alkaline soils of the Alaskan Seward Peninsula, *P. specuicola* to sandstone canyons of the south-western United-States and *P. alcalina* to alkaline meadows of north-eastern Idaho (Fig. 1a). Such restricted and edaphically specialised distributions have been explained by affinity to cool, moist conditions that supposedly provided micro-refugial habitats for these species (Kelso, 1987).

The contrast between the narrow distribution of diploids in clade A (Fig. 1a) and the broad distribution of most diploids in clade D (Fig. 1d-e) raises intriguing questions about the likely explanations of such different ranges for these heterostylous species. Part of the answer might be sought into the stricter edaphic requirements of *P. alcalina*, *P. specuicola* and *P. anvilensis* (Kelso, 1987). Additionally, the four diploid species of clade A currently occur in areas that were ice-free throughout the Pleistocene (Fig. 1a; Ono, 1985; Siegert, 2001). Consequently, these species may represent remnants of a more broadly distributed ancestral diploid that became fragmented during the advancement of Pleistocene glaciers and survived in refugial areas only (Kelso, 1991, 1992), as suggested for other North-American plant groups (e.g. Soltis *et al.*, 1997; Jaramillo-Correa *et al.*, 2004; Godbout *et al.*, 2005). Detailed comparative studies of the ecological characteristics of the diploid species in clades A and D, along with the evaluation of an absolute time scale for their

diversification, would be necessary to gain further insights into the likely origins of the observed distributional patterns.

The tetraploid *P. borealis*, unlike all other polyploids in *Aleuritia*, is heterostylous. This species displays a coastal, amphi-Beringian distribution from Japan to Alaska, expanding eastwards along the Canadian arctic shore, as far as the Mackenzie River delta (Fig. 1a). Morphological similarities, including rhomboid leaves and large flowers forming symmetrical umbels, prompted the suggestion of a close affinity between *P. borealis* and *P. modesta*, with non-overlapping ranges (Kelso, 1991, 1992). However, our maternal phylogeny supports a close relationship between *P. borealis* and the diploid *P. anvilensis* (BS 95%, Fig. 2), in agreement with present-day overlapping distributions (Fig. 1a). Currently, *P. borealis* occupies previously unglaciated areas, but its wider range supports the proposed higher colonising success of polyploids as compared to their diploid progenitors (Stebbins, 1985; Thompson and Lumaret, 1992).

South-American lineage (clade B)

This clade comprises the three available accessions of the South-American octoploid *P. magellanica* (BS 97%, Figs. 2-3), which ranges from Patagonia to Tierra del Fuego and the Falkland Islands (Fig. 1b). Morphologically, *P. magellanica* resembles two heavily farinose polyploids of North-America, namely *P. incana*, with subcapitate inflorescences and flat bracts, and *P. laurentiana*, a robust species with large flowers. Our current data, based on cpDNA markers, do not support the conclusion that a *P. incana*- or *P. laurentiana*-like ancestor might have provided the maternal parent of this polyploid species, as previously suggested (Richards, 2002).

The phylogenetic position of *P. magellanica* remains unresolved (Fig. 2), indicating the need for additional data. The current topology of the strict consensus tree only allows us to suggest that it likely derived from the same common ancestor of clades C and D.

Amphi-Atlantic lineage (clade C)

This clade, comprising six North-American and four European accessions of high ploidy levels (BS 63%, Fig. 2), is connected to the ancestral haplotype that gave origin to both the South-American and the European/North-American lineages (Fig. 3). *Primula incana* (6x) is restricted to inland clay soils of eastern Alaska, central and western Canada, and the Rocky Mountains, extending as far south as Utah and Colorado; *P. laurentiana* (8x) grows on shores and cliffs of north-eastern North-America, expanding to western Newfoundland and the eastern shore of the Hudson Bay; *P. scandinavica* (8x) is endemic to Norway and part of Sweden; and the amphi-Atlantic *P. stricta* (14x) ranges from the eastern Canadian high Arctic, across Greenland, to northern Scandinavia and Russia, as far as Novaya Zemlya (Fig. 1c). The North-American accessions of the amphi-Atlantic

lineage are nested within the European ones (BS 86%, Fig. 2), suggesting that they were derived from European ancestors (see also Figure 3).

High levels of cross-fertility between the European polyploid *P. scandinavica* and the diploid *P. farinosa* (Arnold and Richards, 1998) had prompted the suggestion that the former might be derived from differentiated populations of the latter (Bruun, 1932; Hultgård, 1990, 1993). A similar pattern had been proposed for the origins of *P. incana* and *P. laurentiana* from a diploid *P. mistassinica*-like ancestor on the North-American continent (Kelso, 1991; 1992). While *P. mistassinica* and *P. farinosa* are both core diploid elements in North-America and Europe, respectively (see Discussion on clade D), the cpDNA phylogeny does not support a direct maternal descent of *P. incana* from a *P. mistassinica*-like ancestor nor of *P. scandinavica* from a *P. farinosa*-like ancestor, because the hexaploid and octoploid taxa are included in a complex amphi-Atlantic clade composed entirely of polyploid elements (Figs. 2-3).

As for *P. laurentiana*, our analyses point to independent origins of different populations ascribed to this taxon. Accessions of *P. laurentiana* collected on the Canadian mainland (laurentiana 3-4) are included in an unresolved polytomy with *P. mistassinica* and *P. egaliksensis* (sub-clade d_2, BS 86%, Figs. 2-3). Conversely, populations of *P. laurentiana* from Newfoundland (laurentiana 1-2) form a highly-supported clade with *P. incana* (clade C, BS 95%, Figs. 2-3). The polyphyly of *P. laurentiana* s. l. and the additivity of its ploidy level (8x) with those of *P. mistassinica* (2x) and *P. incana* (6x) may imply that this octoploid derived from independent hybridisation events involving either a *P. mistassinica*- or a *P. incana*-like ancestor as the maternal parent, respectively (Figs. 2-3). Recurrent origins have been demonstrated in other polyploid taxa (Soltis *et al.*, 2003) and are congruent with proposed models for the evolutionary effects of repeated range fragmentation and expansion concomitant with the glacial cycles of the Pleistocene (Stebbins, 1984, 1985).

Based on chromosome number and cross-fertility evidence, two possible pairs of progenitors have been proposed for the 14-ploid *P. stricta*: the North-American *P. incana* (6x) and *P. laurentiana* (8x), and the European *P. scotica* (6x) and *P. scandinavica* (8x; Kelso, 1991; Kelso, 1992; Arnold and Richards, 1998). The current overlapping distributions of *P. stricta* with *P. laurentiana* in North-America and *P. scandinavica* in Europe further indicate a potential role of the two latter species in the origin of the former (Fig. 1c). While all three sampled accessions of *P. stricta* were included in clade C, they did not form a monophyletic group (Fig. 2). The sole North-American accession of *P. stricta* (stricta 3) formed a well-supported sub-clade with the North-American *P. incana* and *P. laurentiana* (laurentiana 1-2, BS 86%), while the European accessions (stricta 1-2) were sister to the rest of clade C (BS 63%), which included one of the proposed

European parents, *P. scandinavica* (Fig. 2). Therefore, different *P. stricta* populations might have independent origins from two different sets of progenitors on the two sides of the North Atlantic, providing yet another example of recurrent polyploid origins (Soltis *et al.*, 2003). Analyses of nuclear markers might allow us to further clarify the origins of *P. stricta*.

European/North-American lineage (clade D)

This clade (BS 98%, Figs. 2-3), comprising both diploid heterostyles and polyploid homostyles, can be subdivided in two geographical groups. Sub-clade d_1 (BS 59%) encompasses exclusively European species, whereas sub-clade d_2 (BS 86%) North-American species (Figs. 2-3).

The European sub-clade d_1 (BS 59%; Figs. 2-3) may be regarded as a polyploid series centred on the diploid *P. farinosa*. *Primula farinosa* displays a disjunct distribution including the Spanish Sierra Nevada and Pyrenees, the Alps, the Carpathians, and the Baltic region (Fig. 1d). *Primula frondosa* (also 2x) is endemic to northeast Bulgaria, *P. halleri* (4x) to the eastern Alps, and *P. scotica* (6x) to northern Scotland and nearby Orkney islands (Fig. 1d). Considering the widespread range and diploid state of *P. farinosa*, this species has been proposed as the "ancestor" of most extant European taxa (Bruun, 1932; Hultgård, 1990, 1993). Different accessions of *P. farinosa* form a polytomy with the clade comprising *P. frondosa*, *P. halleri* and *P. scotica* (BS 63%, Fig. 2), thus not rejecting its proposed central role in Europe. Furthermore, the central position occupied by the "alpine" haplotype of *P. farinosa* (farinosa 1, 3-5) in the corresponding part of the network (Fig. 3) suggests that it represents the ancestral haplotype of clade D (Crandall and Templeton, 1993) and that this lineage diversified from central Europe. The haplotype network further suggests that the North-American lineage d_2 might have derived from an "alpine"-like haplotype of *P. farinosa*.

The cpDNA sequences of *P. farinosa* are characterised by infra-specific polymorphism (Fig. 3). The five detected haplotypes are geographically structured, with two haplotypes in the Alps (farinosa 1, 3-5 and 7), two in the Pyrenees (farinosa 2 and 6), and one in the Carpathians (farinosa 8-9; Fig. 3), suggesting past isolation in separate refugia (Hewitt, 2004). Phylogeographic studies identified similar patterns of geographical structuring in the genetic diversity of *Trollius europaeus* (Després *et al.*, 2002), *Pritzelago alpina* (Kropf *et al.*, 2003) and *Ranunculus glacialis* (Schönswetter *et al.*, 2003), supporting survival during glacial maxima in different southern refugia. Following glacial retreat, Northern Europe was likely re-colonised by more southerly source populations either in the Alps or in the Carpathians. The haplotype shared by alpine (farinosa 1, 3) and British populations (farinosa 4-5) of *P. farinosa* (Fig. 3) denotes a possible migratory route between the central Alps and the British Isles. The investigation of Scandinavian populations of *P. farinosa* is likely to provide further evidence to elucidate post-glacial migration routes between northern and southern Europe.

The diploid *P. frondosa* forms a weakly supported polytomic clade with *P. halleri* (4x) and *P. scotica* (6x; BS 63%, Fig. 2), suggesting that a *P. frondosa*-like ancestor might have been involved in the origin of either polyploid species. An allopolyploid origin for *P. scotica* had been suggested based on the high levels of allelic variation found in its populations (Glover and Abbott, 1995).

The diploid *P. mistassinica* in clade d_2 (BS 86%; Figs. 2-3) may constitute the North-American counterpart of the European *P. farinosa* in clade d_1 (Vogelmann, 1956; Kelso, 1991, 1992). This species is widely distributed across the boreal forest of North-America (Fig. 1e), and forms a diploid-polyploid complex with *P. egaliksensis* (4x) and the Canadian accessions (laurentiana 3-4) of *P. laurentiana* (8x). Similarly to *P. farinosa*, the polymorphism detected in the cpDNA sequences of *P. mistassinica*, with one haplotype to the east (mistassinica 1-6) and two haplotypes to the west of the Great Lakes (mistassinica 7-8 and 9), might reflect the influence of Pleistocene glacial cycles on infra-specific genetic structuring (Fig. 3).

At least five glacial refugia have been identified in North-America: i) east of the Appalachians; ii) along the North-Atlantic coast of Canada (Newfoundland); iii) in the Midwest, south of the Laurentide Ice Sheet; iv) in the west, south of the Cordilleran Ice Sheet; and v) in the ice-free region of Beringia (e.g. Soltis *et al.*, 1997; Siegert, 2001; Jaramillo-Correa *et al.*, 2004; Godbout *et al.*, 2005). More specifically, Vogelmann (1956) proposed the Appalachians, Newfoundland, and the Midwest as potential refugia for *P. mistassinica*. Our results suggest that isolated populations of *P. mistassinica* might have survived in either the Appalachians or Newfoundland, giving origin to haplotype 1-6, while haplotypes 7-8 and 9 might be derived from relictual populations either in the Midwest, the west, or Beringia. Finally, the past presence of vast glacial lakes in central North-America – especially Lake Agassiz between 13,000 and 8,000 years ago – may account for the east-west partition of haplotypes in *P. mistassinica* (Siegert, 2001).

The tetraploid *P. egaliksensis* extends from the Russian Bering Sea coast to North-America, Greenland, and Iceland, where it is likely to be extinct (H. Kristinsson, pers. comm.; Fig. 1e). This species has traditionally been placed in the primarily Asian section *Armerina*, based on similar morphological characteristics, including lack of farina, entire petiolate leaves, and narrow elongate capsules. However, the overlapping distributions of *P. egaliksensis*, *P. mistassinica* and to some extent *P. nutans* (sect. *Armerina*), a chromosome number for *P. egaliksensis* that is additive between diploids of sects. *Aleuritia* and *Armerina*, and gland type, pollen size, colpi number and exine reticulations that are intermediate between sects. *Aleuritia* and *Armerina* suggested that this tetraploid species might represent the product of hybridisation between a *P. mistassinica*- and a *P. nutans*-like ancestor (Kelso, 1991, 1992). Recent phylogenetic analyses based on two chloroplast markers and single accessions indicated that the *P. mistassinica* lineage might have provided the

maternal parent of *P. egaliksensis* (Mast *et al.*, 2001; Mast *et al.*, 2006). The present study, encompassing larger infra-specific sampling and three additional cpDNA markers, does not contradict the former phylogenetic conclusion, for all accessions of *P. egaliksensis* fall in the same clade with samples of *P. mistassinica* (BS 86%, Fig. 2).

Evolution of polyploid, homostylous species in Aleuritia

The phylogenetic analyses performed for the current study inferred a diploid, heterostylous most recent common ancestor for *Aleuritia* and multiple origins for the polyploid, homostylous lineages (Figs. 2, 4). The outcome of concentrated changes tests rejected the null hypothesis that switches in breeding systems are randomly distributed across the phylogeny (Fig. 4), thus providing explicit, novel evidence in favour of the proposed correlation between polyploidy and homostyly in *Aleuritia* (Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). Indeed, an association between these two conditions was also described in Boraginaceae (Schoen *et al.*, 1997), Rubiaceae (Naiki and Nagamasu, 2004), and Turneraceae (Barrett and Shore, 1987; Truyens *et al.*, 2005).

The increased rates of recombination that appear to be typical of some polyploids (Wendel, 2000; Soltis *et al.*, 2003) might favour the disruption of the heterostyly linkage group (Dowrick, 1956; Wedderburn and Richards, 1992), thus providing a possible explanation for the frequent association between polyploidy and homostyly. The hypothesis that the switch from diploidy to polyploidy might favour the subsequent switch to homostyly is further supported by the fact that one *Aleuritia*-species is polyploid and heterostylous (*P. borealis*), whereas none is known to be diploid and homostylous (Table 1). At the same time, polyploidy may have reduced the potentially deleterious effects of inbreeding depression in self-compatible homostylous species, because polyploidisation appears to be associated with an increase of genetic diversity (Lande and Schemske, 1985; Thompson and Lumaret, 1992).

It has been suggested that most *Aleuritia* polyploids are of hybrid origin (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Glover and Abbott, 1995). Studies on synthetic polyploids of *Brassica* showed that the frequency of genomic rearrangements in polyploids may be associated with the degree of genomic divergence between the diploid parents (Song *et al.*, 1995; Wendel, 2000; Soltis *et al.*, 2003). Therefore, allopolyploids, derived from interspecific hybridisation, may experience higher rates of genomic rearrangements than autopolyploids derived from conspecific parents (Stebbins, 1950). Hence, the prevalence of homostyly in polyploid lineages of *Aleuritia* may be primarily linked to hybridisation, rather than chromosome doubling alone (see also Kelso, 1992). In this context, the maintenance of a heteromorphic breeding system in the tetraploid *P. borealis* raises interesting questions concerning

the origin of polyploidy in this species, whether it occurred via allo- or autopolyploidy and whether it occurred recently or not.

In *Aleuritia*, polyploid, self-compatible homostyles are found at high latitudes (Table 1), so that a close association between distribution, ploidy levels, and breeding systems has been proposed (Bruun, 1932; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). However, the diploid, heterostylous *P. anvilensis* and the polyploid, heterostylous *P. borealis* also occupy very high latitudes (Table 1), but in areas that probably remained ice-free throughout the Pleistocene (Fig. 1a; Ono, 1985; Siegert, 2001). In this regard, it is also useful to compare the distribution of heterostylous and homostylous species that now occur, partially or entirely, in previously glaciated areas on the same continent (Siegert, 2001), e.g. *P. mistassinica* (2x) with *P. egaliksensis* (4x), *P. laurentiana* (8x; Fig. 1e) and *P. incana* (6x; Fig. 1c) in North-America, and *P. farinosa* (2x) with *P. scotica* (6x; Fig. 1d), *P. scandinavica* (8x) and *P. stricta* (14x; Fig. 1c) in Europe. This comparison suggests that the higher polyploid homostyles tend to expand further north into previously glaciated areas than the diploid heterostyles (see also Table 1), implying that the switch to autogamy may have represented a selective advantage in the colonisation of barren post-glacial habitats (Stebbins, 1957; Kelso, 1992).

Indeed, heterostylous populations of *Aleuritia* possibly growing at the ice-sheet margins may have suffered from mate deficiency, as predicted by the "abundant centre model", according to which fewer mates are available at the edge than at the centre of a species' range (Hengeveld and Haeck, 1982). The challenges due to mate deficiency might have been compounded by the reduced pollinator activity (McCall and Primack, 1992; Totland, 1994) associated with very low temperatures during the glacial maxima (Siegert, 2001). Furthermore, insect visitation rates appear to be positively correlated with flower density, thus pollinator visits decrease at the edge of a species' range (Thomson, 1981; Jacquemyn *et al.*, 2002). Consequently, heterostylous, self-incompatible individuals, which depend on pollinators for successful fertilisation, may have experienced lower reproductive success at the fringe of the Quaternary ice sheets resulting from mate and pollen limitation. Under such ecological conditions, selection for reproductive assurance may have favoured the establishment of homostylous, self-compatible mutants that were independent of both mate density and pollinator activity (Stebbins, 1957; Fausto *et al.*, 2001; Kalisz *et al.*, 2004).

In general, the wider distribution of homostylous than heterostylous *Aleuritia* species in previously glaciated areas of high latitudes may reflect the higher colonisation potential conferred by autogamy (Baker, 1955; Stebbins, 1957). Indeed, homostylous, self-compatible *Aleuritia* species were probably better adapted to invade habitats opened by glacier retreat, because only a few

individuals would have been sufficient to found a new population (Baker, 1955; Stebbins, 1957). In this regard, it is relevant to note that the most disjunct species of *Aleuritia* (and indeed, of the entire *Primula*), i.e. *P. magellanica*, is a homostylous octoploid (Fig. 1b).

To summarise, the switch to homostyly played an important role in regulating the success of polyploid *Aleuritia* species, by allowing the newly arisen polyploids to self-fertilise and expand their range (Baker, 1955; Stebbins, 1957; Thompson and Lumaret, 1992). Therefore, the high frequency of polyploid, autogamous species in the Arctic may be the product of selection for increased selfing ability in habitats where pollination is unreliable (Molau, 1993).

Acknowledgments

For help with the collection of tissue samples we thank Torbjørn Alm, Alan Batten, Bruce Bennett, Sean Blaney, Christian Brochmann, Luc Brouillet, Edward Buyarski, Matthew Carlson, Mary Beth Cook, Reidar Elven, Pamela Eveleigh, Tupuna Kovanen, Wendy Maovlic, Michael Oldham, Carolyn Parker, Carl Roland, Peter Schönswetter, Heidi Solstad, Andreas Tribsch; all the colleagues listed in Appendix 1; and the following herbaria: ACAD, ALA, ALAJ, ALTA, AMNH, CAFB, CAN, CDFN, DAO, GB, HAM, ICEL, LKHD, MT, NFLD, NHIC, NSPM, O, QFA, S, SASK, SCFQ, SLU, SMI, TRH, TROM, TRTE, S, UBC, UME, UNB, UQTR, UPS, UWPG, WAT, WIN, WLK. We would also like to acknowledge John Richards for general guidance on *Primula*, Péter Szövényi for help in the interpretation of network analyses, and two anonymous reviewers for helpful comments on earlier versions of the manuscript. The Georges und Antoine Claraz-Schenkung and the Swiss Academy of Sciences funded AG's field trips. This study was initially facilitated by grant no. 3100-061674.00/1 from the Swiss National Science Foundation and funds from the Office for Equal Opportunity of the University of Zurich to EC.

References

Abbott, R. J., and C. Brochmann. 2003. History and evolution of the arctic flora: in the footsteps of Eric Hultén. Molecular Ecology 12:299-313.

Arnold, E. S., and A. J. Richards. 1998. On the occurrence of unilateral incompatibility in *Primula* section *Aleuritia* Duby and the origin of *Primula scotica* Hook. Botanical Journal of the Linnean Society 128:359-368.

Baker, H. G. 1955. Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347-349.

Barrett, S. C. H., and J. S. Shore. 1987. Variation and evolution of breeding systems in the *Turnera ulmifolia* L. complex (Turneraceae). Evolution 41:340-354.

Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. Biological Journal of the Linnean Society 82:521-536.

Bruun, H. G. 1932. Cytological studies in *Primula*, with special reference to the relation between the karyology and taxonomy of the genus. Symbolae Botanicae Upsalienses 1:1-239.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.

Conti, E., E. Suring, D. Boyd, J. Jorgensen, J. Grant, and S. Kelso. 2000. Phylogenetic relationships and character evolution in *Primula* L.: the usefulness of ITS sequence data. Plant Biosystems 134:385-392.

Corriveau, J. L., and A. W. Coleman. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperms species. American Journal of Botany 75:1443-1458.

Crandall, K. A. 1994. Intraspecific cladogram estimation: accuracy at higher levels of divergence. Systematic Biology 43:222-235.

Crandall, K. A., and A. R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959-969.

DeBry, R. W., and R. G. Olmstead. 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. Systematic Biology 49:171-179.

Després, L., S. Loriot, and M. Gaudeul. 2002. Geographic pattern of genetic variation in the European globeflower *Trollius europaeus* L. (Ranunuculaceae) inferred from amplified fragment length polymorphism markers. Molecular Ecology 11:2337-2347.

Dowrick, V. P. J. 1956. Heterostyly and homostyly in *Primula obconica*. Heredity 10:219-236.

Dynesius, M., and R. Jansson. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proceedings of the National Academy of Science, USA 97:9115-9120.

Fausto, J. A., V. M. Eckhart, and M. A. Geber. 2001. Reproductive assurance and the evolutionary ecology of self-pollination in *Clarkia xantiana* (Onagraceae). American Journal of Botany 88:1794-1800.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the boostrap. Evolution 17:223-263.

Glover, B. J., and R. J. Abbott. 1995. Low genetic diversity in the Scottish endemic *Primula scotica* Hook. New Phytologist 129:147-153.

Godbout, J., J. P. Jaramillo-Correa, J. Beaulieu, and J. Bousquet. 2005. A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. Molecular Ecology 14:3497-3512.

Grundt, H. H., M. Popp, C. Brochmann, and B. Oxelman. 2004. Polyploid origins in a circumpolar complex in *Draba* (Brassicaceae) inferred from cloned nuclear DNA sequences and fingerprints. Molecular Phylogenetics and Evolution 32:695-710.

Hengeveld, R., and J. Haeck. 1982. The distribution of abundance. I. Measurements. Journal of Biogeography 9:303-316.

Hewitt, G. H. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 359:183-195.

Hultgård, U.-M. 1990. Polyploidy and differentiation in N European populations of *Primula* subgenus *Aleuritia*. Sommerfeltia 11:117-135.

Hultgård, U.-M. 1993. *Primula scandinavica* and *P. stricta* - patterns of distribution, variation, reproductive strategies and migrations. Opera Botanica 121:35-43.

Jacquemyn, H., R. Brys, and M. Hermy. 2002. Patch occupancy, population size and reproductive success of a forest herb (*Primula elatior*) in a fragmented landscape. Oecologia 130:617-625.

Jaramillo-Correa, J. P., J. Beaulieu, and J. Bousquet. 2004. Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American conifer. Molecular Ecology 13:2735-2747.

Kalisz, S., D. W. Vogler, and K. M. Hanley. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. Nature 430:884-887.

Kelso, S. 1987. *Primula anvilensis* (Primulaceae): a new species from northwestern Alaska. Systematic Botany 12:9-13.

Kelso, S. 1991. Taxonomy of *Primula* sects. *Aleuritia* and *Armerina* in North America. Rhodora 93:67-99.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. Arctic and Alpine Research 24:82-87.

Kohn, J. R., S. W. Graham, B. Morton, J. J. Doyle, and S. C. H. Barrett. 1996. Reconstruction of the evolution of reproductive characters in Pontederiaceae using phylogenetic evidence from chroroplast DNA restriction-site variation. Evolution 50:454-469.

Kropf, M., J. W. Kadereit, and H. P. Comes. 2003. Differential cycles of range contraction and expansion in European high mountain plants during the late Quaternary: insights from *Pritzelago alpina* (L.) O. Kuntze (Brassicaceae). Molecular Ecology 12:931-949.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24-40.

Leitch, I. J., and M. D. Bennett. 2004. Genome downsizing in polyploid plants. Biological Journal of the Linnean Society 82:651-663.

Lihová, J., J. Fuertes Aguilar, K. Marhold, and G. Nieto Feliner. 2004. Origin of the disjunct tetraploid *Cardamine amporitana* (Brassicaceae) assessed with nuclear and chloroplast DNA sequence data. American Journal of Botany 91:1231-1242.

Löve, Á., and D. Löve. 1975. Cytotaxonomical Atlas of the Arctic Flora. J. Cramer, Vaduz, Liechtenstein.

Maddison, W. P. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? Evolution 44:539-557.

Maddison, W. P., and D. R. Maddison. 2000. MacClade, version 4. Sinauer Associates.

Mast, A. R., S. Kelso, and E. Conti. 2006. Are any primroses (*Primula*) primitively monomorphic? New Phytologist 171:605-616.

Mast, A. R., S. Kelso, A. J. Richards, D. J. Lang, D. M. S. Feller, and E. Conti. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. International Journal of Plant Sciences 162:1381-1400.

McCall, C., and R. B. Primack. 1992. Influence of flower characteristics, weather, time of day, and season on insect visitation rates in three plant communities. American Journal of Botany 79:434-442.

Molau, U. 1993. Relationships between flowering phenology and life history strategies in tundra plants. Arctic and Alpine Research 25:391-402.

Mort, M. E., P. S. Soltis, D. E. Soltis, and M. L. Mabry. 2000. Comparison of three methods for estimating internal support on phylogenetic trees. Systematic Biology 49:160-171.

Naiki, A., and H. Nagamasu. 2004. Correlation between distyly and ploidy level in *Damnacanthus* (Rubiaceae). American Journal of Botany 91:664-671.

Ono, Y. 1985. Late glacial paleoclimate reconstructed from glacial and periglacial landforms in Japan. Geographical Review of Japan, Series B 57:80-100 (in Japanese).

Oxelman, B., M. Liden, and D. Berglund. 1997. Chloroplast *rps*16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). Plant Systematics and Evolution 206:393-410.

Posada, D., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. Trends in Ecology and Evolution 16:37-45.

Ramsey, J., and D. W. Schemske. 2002. Neopolyploidy in flowering plants. Annual Review of Ecology and Systematics 33:589-639.

Richards, J. 2002. *Primula*. B. T. Batsford, London, UK.

Scheen, A.-C., C. Brochmann, A. K. Brysting, R. Elven, A. Morris, D. E. Soltis, P. S. Soltis, and V. A. Albert. 2004. Northern hemisphere biogeography of *Cerastium* (Caryophyllaceae): insights from phylogenetic analysis of noncoding plastid nucleotide sequences. American Journal of Botany 91:943-952.

Schoen, D. J., M. O. Johnston, A.-M. L'Heureux, and J. V. Marsolais. 1997. Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). Evolution 51:1090-1099.

Schönswetter, P., O. Paun, A. Tribsch, and H. Niklfeld. 2003. Out of the Alps: colonization of Northern Europe by East Alpine populations of the glacier buttercup *Ranunculus glacialis* L. (Ranunculaceae). Molecular Ecology 12:3373-3381.

Siegert, M. J. 2001. Ice Sheets and Late Quaternary Environmental Change. Wiley, Chichester, UK.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. American Journal of Botany 85:1301-1315.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2003. Advances in the study of polyploidy since *Plant Speciation*. New Phytologist 161:173-191.

Soltis, D. E., M. A. Gitzendanner, D. D. Strenge, and P. S. Soltis. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Systematics and Evolution 206:353-373.

Song, K., P. Lu, K. Tang, and T. C. Osborn. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Science, USA 92:7719-7723.

Stebbins, G. L. 1950. Variation and Evolution in Plants. Columbia University Press, New York, NY, USA.

Stebbins, G. L. 1957. Self-fertilization and population variability in the higher plants. The American Naturalist 91:337-354.

Stebbins, G. L. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. Botanica Helvetica 94:1-13.

Stebbins, G. L. 1985. Polyploidy, hybridization, and the invasion of new habitats. Annals of the Missouri Botanical Garden 72:824-832.

Swofford, D. L. 1999. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), version 4. Sinauer Associates.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17:1105-1109.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619-633.

Thompson, J. D., and R. Lumaret. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. Trends in Ecology and Evolution 7:302-307.

Thomson, J. D. 1981. Spatial and temporal components of resource assessment by flower-feeding insects. Journal of Animal Ecology 50:49-59.

Totland, Ø. 1994. Influence of climate, time of day and season, and flower density on insect flower visitation in alpine Norway. Arctic and Alpine Research 26:66-71.

Trift, I., M. Källersjö, and A. A. Anderberg. 2002. The monophyly of *Primula* (Primulaceae) evaluated by analysis of sequences from the chloroplast gene *rbc*L. Systematic Botany 27:396-407.

Truyens, S., M. M. Arbo, and J. S. Shore. 2005. Phylogenetic relationships, chromosome and breeding system evolution in *Turnera* (Turneraceae): inferences from its sequence data. American Journal of Botany 92:1749-1758.

Verheyen, E., W. Salzburger, J. Snoeks, and A. Meyer. 2003. Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. Science 300:325-329.

Vogelmann, H. W. 1956. A Biosystematic Study of *Primula mistassinica* Michx. University of Michigan, Ann Arbor, MI, USA.

Wedderburn, F., and A. J. Richards. 1992. Secondary homostyly in *Primula* L.: evidence for the model of the 'S' supergene. New Phytologist 121:649-653.

Wendel, J. F. 2000. Genome evolution in polyploids. Plant Molecular Biology 42:225-249.

Young, N. D., and J. Healy. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4.

Appendix 1. Species of *Primula* sampled for the present study. Data are presented in the following sequence: Taxon, Section, OTU abbreviation (in bold), Population location, Voucher information (italicised), GenBank accession numbers for *rpl*16 intron, *rps*16 intron, *trn*L intron, *trn*L-F spacer and *trn*T-L spacer, respectively. GenBank accession numbers with literature references are for sequences obtained by other authors. na, not available.

Primula alcalina Cholewa & Henderson, sect. *Aleuritia* subsect. *Aleuritia*, **alcalina**, USA, Idaho, Targhee National Forest, *Lehman s.n.* (Z), AF402489 (Mast *et al.*, 2001), DQ379908, AF402370 (Mast *et al.*, 2001), DQ379731, DQ379792

Primula anvilensis Kelso, sect. *Aleuritia* subsect. *Aleuritia*, **anvilensis**, USA, Alaska, Nome, *Guggisberg & Mansion 200703-1* (Z), DQ379852, DQ379911, DQ379733, DQ379733, DQ379794

Primula borealis Duby, sect. *Aleuritia* subsect. *Aleuritia*, **borealis1**, USA, Alaska, Unalakleet, *Parker 7537* (ALA), AF402488 (Mast *et al.*, 2001), DQ379913, AF402369 (Mast *et al.*, 2001), DQ379735, DQ379796; **borealis2**, USA, Alaska, Kotzebue, *Elven & Solstad S03-11* (O), DQ379853, DQ379912, DQ379736, DQ379736, DQ379797

Primula egaliksensis Wormsk. ex Hornem., sect. *Armerina*, **egaliksensis1**, Canada, Yukon Territory, Abraham Mt., *Elven & Solstad SUP03-352* (O), DQ379875, DQ379936, DQ379759, DQ379759, DQ379817; **egaliksensis2**, Canada, Newfoundland, St.-Barbe, *Guggisberg & Mansion 250604-8* (Z), DQ379876, DQ379937, DQ379763, DQ379763, DQ379823; **egaliksensis3**, USA, Alaska, Goodnews Bay, *Parker 160604-2* (Z), DQ379877, DQ379940, DQ379762, DQ379762, DQ379826; **egaliksensis4**, USA, Colorado, South Park, *Kelso 060703-2* (Z), DQ379878, DQ379938, DQ379764, DQ379764, DQ379824; **egaliksensis5**, USA, Alaska, Glacier Bay NP, Gustavus, *Bosworth 160603-1* (Z), DQ379879, DQ379939, DQ379765, DQ379765, DQ379825; **egaliksensis6**, USA, Alaska, Wrangell-St.Elias NP, Jake Lake, *Guggisberg & Mansion 230603-1* (Z), DQ379880, DQ379941, DQ379761, DQ379761, DQ379827; **egaliksensis7**, Canada, Newfoundland, Port au Choix Peninsula, *Guggisberg & Mansion 240604-6b* (Z), DQ379881, DQ379942, DQ379760, DQ379760, DQ379828; **egaliksensis8**, USA, Alaska, Cape Krusenstern, *Parker 10392* (ALA), DQ379882, DQ379943, DQ379766, DQ379766, DQ379829

Primula fasciculata Balf. f. & Ward, sect. *Armerina*, **fasciculata**, horticultural origin, *McBeath s.n.* (Z), DQ379891, DQ379953, DQ379776, DQ379776, DQ379839

Primula farinosa L., sect. *Aleuritia* subsect. *Aleuritia*, **farinosa1**, Italy, South Tirol, Schlüsseljochs, *Schönswetter & Tribsch 100703-3* (Z), DQ379845, DQ379897, DQ379720, DQ379720, DQ379781; **farinosa2**, Spain, Aragon, Central Pyrenees, Macizo La Maladeta, *Schneeweiss & Schönswetter 230703-5* (Z), DQ379846, DQ379898, DQ379722, DQ379722, DQ379784; **farinosa3**, Switzerland, Bern, Faulhornstock, *Wiedermann 230803-1* (Z), DQ379848, DQ379899, DQ379723, DQ379723, DQ379785; **farinosa4**, UK, Teesdale, Sand Syke, *Richards s.n.* (Z), DQ379849, DQ379900, DQ379716, DQ379716, DQ379787; **farinosa5**, UK , Cumbria, Gait Barrow, *Richards s.n.* (Z), AF402474 (Mast *et al.*, 2001), DQ379901, AF402356 (Mast *et al.*, 2001), DQ379725, DQ379786; **farinosa6**, France, Aquitaine, Atlantic Pyrenees, El Portalet, *Schneeweiss & Schönswetter 250703-1* (Z), DQ379842, DQ379893, DQ379724, DQ379724, DQ379782; **farinosa7**, Italy, Cuneo, Cottian Alps, Col de Sampeyre, *Wiedermann 160703-1* (Z), DQ379843, DQ379894, DQ379721, DQ379721, DQ379783; **farinosa8**, Slowakia, W Carpathians, Vel'ka Fatra Mts., *Ronikier & Ronikier 020503-1* (Z), DQ379844, DQ379895, DQ379718, DQ379718, DQ379779; **farinosa9**, Poland, W Carpathians, Beskid Sadecki massif, *Ronikier & Ronikier 120903-1* (na), DQ379847, DQ379896, DQ379719, DQ379719, DQ379780

Primula frondosa Janka, sect. *Aleuritia* subsect. *Aleuritia*, **frondosa**, horticultural origin, *Kelso s.n*. (Z), AF402477 (Mast *et al.*, 2001), DQ379902, AF402359 (Mast *et al.*, 2001), DQ379717, DQ379788

Primula halleri Gmel., sect. *Aleuritia* subsect. *Aleuritia*, **halleri**, Austria, Hohe Tauern, Ankogelgruppe, *Schönswetter & Tribsch 310503-1* (Z), DQ379850, DQ379903, DQ379726, DQ379726, DQ379789

Primula incana M. E. Jones, sect. *Aleuritia* subsect. *Aleuritia*, **incana1**, Canada, Yukon Territory, Whitehorse, Takhini Salt Flats, *Guggisberg & Mansion 020703-1* (Z), DQ379855, DQ379915, DQ379742, DQ379742, DQ379799; **incana2**, USA, Alaska, Fairbanks, *Guggisberg & Mansion 080703-1-1* (Z), DQ379854, DQ379914, DQ379740, DQ379740, DQ379798; **incana3**, Canada, Yukon Territory, Meadow Lake, *Guggisberg & Mansion 040703-3* (Z), DQ379856, DQ379916, DQ379741, DQ379741, DQ379800

Primula involucrata Wallich, sect. *Armerina*, **involucrata**, horticultural origin, *Carrie s.n.* (Z), DQ379892, DQ379954, DQ379777, DQ379777, DQ379840

Primula ioessa Smith, sect. *Sikkimensis*, **ioessa**, horticultural origin, *Tromsoy Botanical Garden 92-1052* (na), AF402501 (Mast *et al.*, 2001), DQ379955, AF402381 (Mast *et al.*, 2001), DQ379778, DQ379841

Primula laurentiana Fernald, sect. *Aleuritia* subsect. *Aleuritia*, **laurentiana1**, Canada, Newfoundland, Green Point, *Guggisberg & Mansion 230604-3* (Z), DQ379857, DQ379917, DQ379744, DQ379744, DQ379803; **laurentiana2**, Canada, Newfoundland, Reefs Harbour, *Guggisberg & Mansion 260604-1* (Z), DQ379858, DQ379918, DQ379743, DQ379743, DQ379804; **laurentiana3**, Canada, Quebec, St.-Lawrence River, St.-Simon, *Guggisberg & Mansion 130604-2* (Z), DQ379859, DQ379920, DQ379747, DQ379747, DQ379802; **laurentiana4**, Canada, New Brunswick, Fundy Bay, *Guggisberg & Mansion 180604-1* (Z), DQ379860, DQ379919, DQ379746, DQ379746, DQ379801

Primula magellanica Lehm., sect. *Aleuritia* subsect. *Aleuritia*, **magellanica1**, Argentina, Tierra del Fuego, *Richards s.n.* (na), DQ379865, DQ379922, DQ379739, DQ379739, DQ379812; **magellanica2**, E Falkland Is., Stanley aiport, *Thompson & Woods 251103-1* (na), DQ379864, DQ379921, DQ379737, DQ379737, DQ379810; **magellanica3**, horticultural origin, *Lever s.n.* (na), AF402482 (Mast *et al.*, 2001), DQ379923, AF402364 (Mast *et al.*, 2001), DQ379738, DQ379811

Primula misstassinica Michaux, sect. *Aleuritia* subsect. *Aleuritia*, **misstassinica1**, Canada, Quebec, St.-Lawrence River, Beaumont, *Guggisberg & Mansion 130604-1* (Z), DQ379869, DQ379931, DQ379751, DQ379751, DQ379818; **misstassinica2**, Canada, Newfoundland, Fischells Brook, *Guggisberg & Mansion 210604-1* (Z), DQ379870, DQ379927, DQ379753, DQ379753, DQ379813; **misstassinica3**, Canada, Ontario, Lake Huron, Manitoulin I., *Guggisberg & Mansion 100604-2* (Z), DQ379871, DQ379930, DQ379758, DQ379758, DQ379816; **misstassinica4**, Canada, Ontario, Lake Huron, Port Elgin, *Guggisberg & Mansion 090604-1* (Z), DQ379872, DQ379929, DQ379757, DQ379757, DQ379815; **misstassinica5**, Canada, Newfoundland, Reefs Harbour, *Guggisberg & Mansion 260604-1* (Z), DQ379873, DQ379934, DQ379754, DQ379754, DQ379821; **misstassinica6**, Canada, New Brunswick, Rothesay, *Guggisberg & Mansion 170604-2* (Z), DQ379874, DQ379932, DQ379752, DQ379752, DQ379819; **mistassinica7**, Canada, British Columbia, Boya Lake, *Guggisberg & Mansion 300603-1* (Z), DQ379866, DQ379928, DQ379755, DQ379755, DQ379814; **mistassinica8**, USA, Wisconsin, Newport State Park, *Anderson 210503-1a* (Z), DQ379867, DQ379933, DQ379750, DQ379750, DQ379820; **mistassinica9**, Canada, Alberta, Bow Valley; *Eveleigh 300603-3* (Z), DQ379868, DQ379935, DQ379756, DQ379756, DQ379822

Primula modesta Bisset & Moore, sect. *Aleuritia* subsect. *Aleuritia*, **modesta**, horticultural origin, *Richards s.n.* (Z), AF402490 (Mast *et al.*, 2001), DQ379910, AF402371 (Mast *et al.*, 2001), DQ379734, DQ379795

Primula nutans Georgi, sect. *Armerina*, **nutans1**, Mongolia, Muren Hovsgol aimag, Delgermuren River valley, *Oyumaa 090703-1* (Z), DQ379883, DQ379944, DQ379767, DQ379767, DQ379830; **nutans2**, Russia, Altai Republic, valley of Tarkhata River, *Tribsch 300703-1* (Z), DQ379886, DQ379949, DQ379773, DQ379773, DQ379834; **nutans3**, Mongolia, Muren Hovsgol aimag, Hatgal Huzuuvch, *Oyumaa 180703-1* (Z), DQ379889, DQ379950, DQ379772, DQ379772, DQ379836; **nutans4**, Russia, Irkutsk region, Lake Baikal, *Kovtonyuk s.n.* (na), DQ379890, DQ379951, DQ379774, DQ379774, DQ379838; **nutans5**, USA, Alaska, Northway, *Guggisberg & Mansion 240603-2* (Z), DQ379884, DQ379946, DQ379769, DQ379769, DQ379831; **nutans6**, USA, Alaska, Nome, *Guggisberg & Mansion 220703-6* (Z), DQ379885, DQ379945, DQ379768, DQ379768, DQ379832; **nutans7**, Canada, Yukon Territory, Little Braeburn Lake, *Guggisberg & Mansion 030703-2* (Z), DQ379887, DQ379947, DQ379770, DQ379770, DQ379833; **nutans8**, USA, Alaska, Goodnews Bay, *Parker 150604-5* (Z), DQ379888, DQ379948, DQ379771, DQ379771, DQ379835; **nutans9**, USA, Alaska, Unalakleet, *Parker 6966* (ALA), AF402494 (Mast *et al.*, 2001), DQ379952, AF 402374 (Mast *et al.*, 2001), DQ379775, DQ379837

Primula scandinavica (Bruun) Bruun, sect. *Aleuritia* subsect. *Aleuritia*, **scandinavica1**, Norway, Tromsoy, *Alsos & Tveraabak 250803-1* (Z), DQ379861, DQ379906, DQ379729, DQ379729, DQ379805; **scandinavica2**, Norway, Bodo, *Richards s.n.* (Z), AF402483 (Mast *et al.*, 2001), DQ379907, AF402365 (Mast *et al.*, 2001), DQ379730, DQ379806

Primula scotica Hooker, sect. *Aleuritia* subsect. *Aleuritia*, **scotica1**, horticultural origin, *Tromsoy Botanical Garden s.n.* (na), DQ379851, DQ379904, DQ379727, DQ379727, DQ379791; **scotica2**, horticultural origin, *Lever s.n.* (na), AF402476 (Mast *et al.*, 2001), DQ379905, AF402358 (Mast *et al.*, 2001), DQ379728, DQ379790

Primula specuicola Rydb., sect. *Aleuritia* subsect. *Aleuritia*, **specuicola**, USA, Utah, Negro Bill Canyon, *Kelso 00-59* (COCO), AF402487 (Mast *et al.*, 2001), DQ379909, AF402368 (Mast *et al.*, 2001), DQ379732, DQ379793

Primula stricta Hornem., sect. *Aleuritia* subsect. *Aleuritia*, **stricta1**, horticultural origin, *Tromsoy Botanical Garden s.n.* (na), DQ379863, DQ379925, DQ379748, DQ379748, DQ379808; **stricta2**, horticultural origin, *Mattingley s.n.* (na), AF402484 (Mast *et al.*, 2001), DQ379926, AF402366 (Mast *et al.*,

2001), DQ379749, DQ379809; **stricta3**, Canada, Newfoundland, L'Anse-aux-Meadows, *Guggisberg & Mansion 250604-2* (Z), DQ379862, DQ379924, DQ379745, DQ379745, DQ379807

Chapter 3

Disentangling reticulate evolution in an arctic-alpine polyploid complex

with G. Mansion and E. Conti

Systematic Biology (tentatively accepted)

Abstract

Although polyploidy plays a fundamental role in plant evolution, the elucidation of polyploid origins is fraught with methodological challenges. For example, allopolyploid species may confound phylogenetic reconstruction, because commonly used methods are designed to trace divergent, rather than reticulate patterns. Recently developed techniques of phylogenetic network estimation allow for a more effective identification of incongruence among trees. However, incongruence can also be caused by incomplete lineage sorting, paralogy, concerted evolution and recombination, thus initial hypotheses of hybridisation need to be examined via additional sources of evidence, including the partitioning of infra-specific genetic polymorphisms, morphological characteristics, chromosome numbers, crossing experiments, and distributional patterns. *Primula* sect. *Aleuritia* subsect. *Aleuritia* (*Aleuritia*) represents an ideal case study to examine reticulation because specific hypotheses have been derived from morphology, caryology, inter-fertility and distribution to explain the observed variation of ploidy levels, ranging from diploidy to 14-ploidy. Sequences from five chloroplast and one nuclear ribosomal DNA (nrDNA) markers were analysed to generate the respective phylogenies and consensus networks. Furthermore, extensive cloning of the nrDNA marker allowed for the identification of additive polymorphic sites, investigation of infra-specific genetic polymorphisms via principal coordinate analyses, and detection of recombination between putative progenitor sequences. The results suggest that most polyploids originated via hybridisation and that two taxonomic species formed recurrently from different progenitors, findings that are congruent with the expectations of speciation via secondary contact. Overall, the study highlights the importance of using multiple experimental and analytical approaches to disentangle complex patterns of reticulation.

Introduction

Polyploidisation is one of the most important mechanism of speciation in plants (Soltis and Soltis, 1993; Bretagnolle *et al.*, 1998; Otto and Whitton, 2000; Levin, 2002; Soltis *et al.*, 2003). Indeed, up to 70% of all angiosperms may be of polyploid origin (Stebbins, 1950; Grant, 1981; Masterson, 1994; Otto and Whitton, 2000), including species of small genome size and chromosome number, e.g. *Arabidopsis thaliana* (Vision *et al.*, 2000; Henry *et al.*, 2006). Two types of polyploids are commonly recognised, according to the degree of homology among coexisting genomes. Autopolyploids, which arise within a single species, contain more than two sets of homologous chromosomes in their nuclear genome, while allopolyploids contain more than two sets of homeologous chromosomes that diverged from each other prior to interspecific hybridisation (Ramsey and Schemske, 1998). Importantly, interspecific hybridisation may also give origin to homoploid hybrids that maintain the same ploidy level as their parents (Arnold, 1997; Rieseberg, 1997; Rieseberg and Carney, 1998; Arnold, 2006). Finally, recent evidence suggests that recurrent polyploidisation between the same species often generates polyploids of independent origins that may differ from each other either morphologically and/or molecularly, adding to the complexity of reticulate histories (Soltis and Soltis, 1993; Brochmann *et al.*, 1998; Soltis and Soltis, 1999; Steen *et al.*, 2000).

Species of hybrid origin, either at the homoploid or polyploid level, pose special problems in phylogenetic reconstruction, because they derive from the merging of different evolutionary lineages, while phylogenetic methods are designed to trace relationships among divergent branches of a genealogy (Rieseberg and Ellstrand, 1993; McDade, 1995; Linder and Rieseberg, 2004; Vriesendorp and Bakker, 2005; Arnold, 2006). Morphologically, hybrids vary considerably, because they may express derived, intermediate, or new character states when compared to the parents (Rieseberg and Ellstrand, 1993; McDade, 1995). Hybrids usually inherit the nuclear DNA from both parents and the organellar DNA of only one parent, thus the incongruent placement of individual taxa in phylogenies generated from differently inherited markers can be instrumental in the identification of hypothetical hybrids and their putative parents (Rieseberg and Ellstrand, 1993; Wendel and Doyle, 1998; Linder and Rieseberg, 2004). In angiosperms, comparisons between phylogenies derived from maternally and biparentally inherited chloroplast and nuclear DNA markers, respectively, have been used to reconstruct patterns of reticulation (e.g. Soltis and Kuzoff, 1995; Sang *et al.*, 1997).

Recently developed techniques of phylogenetic network estimation provide a powerful new tool for the identification of potential hybrid origins, because they allow for a more effective visualisation of complex patterns of incongruence among trees, as compared to commonly used

phylogenetic methods (Posada and Crandall, 2001; Linder and Rieseberg, 2004; Morrison, 2005; Huson and Bryant, 2006; Brysting *et al.*, 2007). Importantly, tree incongruence can also be caused by paralogy, incomplete lineage sorting, concerted evolution and recombination (Doyle, 1992; Maddison, 1997; Doyle and Davis, 1998; Wendel and Doyle, 1998; Álvarez and Wendel, 2003; Bailey *et al.*, 2003; Linder and Rieseberg, 2004). Therefore, the inconsistent placement of taxa among trees can only provide an initial hypothesis of hybridisation, which must be further examined by integrating additional sources of evidence, including the partitioning of infra-specific genetic polymorphisms, chromosome number, distributional patterns, morphological characteristics, and inter-fertility experiments (Wendel and Doyle, 1998; Arnold, 2006).

Phylogenies generated from nuclear markers alone can allow for the identification of hybrids and their parents if both parental copies persist in the hybrid genome and if they were sufficiently diverged prior to hybridisation (Doyle and Davis, 1998). In this case, the detection, preferably by cloning, of intra-individual polymorphisms that are additive between the parental nucleotides at the same site provides an indication of possible hybrid origins (Hughes *et al.*, 2002; Smedmark *et al.*, 2003; Devos *et al.*, 2006; Guggisberg *et al.*, 2006a; Noyes, 2006). However, both intra- and interlocus recombination between repeats in multigene families can disrupt patterns of additivity, causing different copies to evolve in concert and potentially leading to complete homogenisation of all repeat types toward one of the parental copies, with the concomitant loss of the other (Wendel *et al.*, 1995; Kovarík *et al.*, 2004; Kovarík *et al.*, 2005). If homogenisation has erased the hybrid signature in the nuclear marker under study, comparison with phylogenies derived from uniparentally inherited markers can still allow for the identification of reticulate patterns, when taxa show incongruent placement in the two kinds of phylogenies (Wendel and Doyle, 1998; Linder and Rieseberg, 2004).

The nuclear marker most commonly used for phylogeny reconstruction at low taxonomic levels has been the internal transcribed spacer (ITS) region of the nuclear ribosomal DNA (nrDNA) cistron, because the high number of repeats and availability of universal primers in the highly conserved flanking regions have greatly facilitated the procedure of *in vitro* amplification in a broad range of distantly related taxa (Baldwin *et al.*, 1995; Álvarez and Wendel, 2003; Small *et al.*, 2004; Nieto Feliner and Rosselló, 2007). The nrDNA cistron, formed by the 18S ribosomal gene, the ITS1, the 5.8S ribosomal gene, the ITS2, and the 26S ribosomal gene, is present in hundreds of copies distributed over one or more loci (Long and Dawid, 1980). ITS polymorphisms have been instrumental in detecting patterns of reticulation in a large number of angiosperms (e.g. Sang *et al.*, 1995; Campbell *et al.*, 1997; Hughes *et al.*, 2002; Fuertes Aguilar and Nieto Feliner, 2003; Devos *et al.*, 2005; Mansion *et al.*, 2005; Devos *et al.*, 2006; Noyes, 2006).

In the present study, we analyse newly sequenced chloroplast DNA and ITS fragments of *Primula* sect. *Aleuritia* subsect. *Aleuritia* Duby (Primulaceae Vent., – hereafter called *Aleuritia*) with a combination of commonly used phylogenetic methods aimed at recovering divergent genealogies (i.e. maximum parsimony and Bayesian inference) and methods recently developed to identify incongruence among trees (i.e. consensus networks; Holland and Moulton, 2003; Holland *et al.*, 2005). In addition, extensive cloning of ITS allowed us to investigate infra-specific genetic polymorphisms via principal coordinate analyses, identify recombination between progenitor sequences, and detect additive polymorphic sites.

The circumboreal, polyploid *Aleuritia* complex is an ideal group with which to study reticulate patterns, because speciation in *Aleuritia* is supposed to be driven by hybridisation. Indeed, ploidy levels in *Aleuritia* range from diploidy to 14-ploidy (Table 1), while morphological, caryological and distributional data suggest hybrid origins for most polyploids, allowing for the identification of putative parents (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993). Furthermore, *Aleuritia* displays an interesting variation of breeding systems, alternating between self-incompatible heterostyly and self-compatible homostyly (Richards, 2002), traits that, in combination with changes of ploidy level, apparently played a fundamental role in the reticulate evolution of the group (Kelso, 1991, 1992; Guggisberg *et al.*, 2006b).

Speciation by hybridisation in *Aleuritia* has been explained within the context of the secondary contact model (Stebbins, 1984, 1985). According to this model, glacial advancement during the Pleistocene caused the fragmentation of diploid, allogamous populations, which may have survived in ice-free areas (i.e. refugia). As glaciers retreated, the differentiated diploid populations came into contact again and hybridised, giving origin to autogamous, polyploid taxa (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). The higher success of the self-compatible, polyploid species at recolonising habitats freed by glacial retreat might be explained in terms of selection for reproductive assurance (Guggisberg *et al.*, 2006b; Carlson *et al.*, 2008). Indeed, all eight diploid species are heterostylous, whereas nine of the ten polyploid species are homostylous, the exception being the tetraploid *P. borealis*. Thus, *Aleuritia* represents an example for the crucial role of allopolyploidy in generating species diversity in the Arctic (Brochmann *et al.*, 2004).

Recently supported as monophyletic in chloroplast-based phylogenies (Mast *et al.*, 2001; Guggisberg *et al.*, 2006b; Mast *et al.*, 2006), *Aleuritia* includes 21 arctic-alpine species of small to medium-sized plants, usually perennial and characterised by a single umbel of flowers, a yellow annulus at the centre of the corolla, dense farina on the stem and calyx, syncolpate pollen, and a base chromosome number of x = 9 (Richards, 2002). The main centres of diversity in *Aleuritia* are

located in the major mountain systems of North-America and Eurasia, except for the Himalayas, and the plains of North-America (Table 1; Richards, 2002). The only South-American species of *Primula* described so far, *P. magellanica*, also belongs to *Aleuritia* (Table 1; Richards, 2002).

Table 1. Distributional ranges and ploidy levels of the 21 species assigned to *Primula* sect. *Aleuritia* subsect. *Aleuritia* (*Aleuritia*) and the two species of sect. *Armerina* (*P. egaliksensis* and *P. nutans*) supposed to form a hybrid complex with *P. mistassinica* (*Aleuritia*). na, not available.

***Primula* species**	**Distributional range**	**Ploidy level**
P. alcalina† Cholewa & Henderson	NE Idaho	2n = 2x = 18
P. anvilensis† Kelso	Seward Penins. (Alaska)	2n = 2x = 18
P. baldschuanica B. Fedstch.	Tadzhikistan, E Afghanistan	na
P. borealis† Duby	NW Alaska, N Canada, NE Siberia	2n = 4x = 36
P. capitellata Boiss.	Iran, Afghanistan, Pakistan	na
P. egaliksensis† Wormsk. in Hornem.	N Iceland, W Greenland, N Canada, Alaska, Chukotsk Penins. (NE Siberia), a few sites in Colorado and NW Wyoming	2n = 4x = 40
P. exigua Velenovsky	SW Bulgaria	2n = 2x = 18
P. farinosa† L.	England, Denmark, Sweden to 64°N, Finnish archipelago, Baltic states to 60°N, Montes Universales (Spain), Pyrenees, Alps, Tatra	2n = 2x = 18
P. frondosa Janka	Stara Planina (SE Bulgaria)	2n = 2x = 18
P. halleri† Gmel.	E Alps, Tatra, Carpathians, ex-Yugoslavia, Albania, Rila and Pirin ranges (Bulgaria)	2n = 4x = 36
P. incana† M. E. Jones	NW America	2n = 6x = 54
P. laurentiana† Fernald	NE America	2n = 8x = 72
P. magellanica† Lehm.	Tierra del Fuego, Patagonia, Falkland Is.	2n = 8x = 72
P. mistassinica† Michaux	Canada, central Alaska	2n = 2x = 18
P. modesta† Bisset & Moore	Japan, Kurile Is.	2n = 2x = 18
P. nutans† Georgi	Gulf of Bothnia, N Fennoscandia eastward to NE Siberia, Kamchatka, Seward Penins. (Alaska), SW Yukon, N Altai, Baikal L., N Mongolia, Gansu Mtns. (China), NW Himalaya	2n = 2x = 22
P. sachalinensis Nakai	E Siberia, Kamchatka and Sakhalin I.	2n = 4x = 36
P. scandinavica† (Bruun) Bruun	Norway, a few sites in Sweden	2n = 8x = 72
P. schlagintweitiana Pax	N Pakistan, Kashmir and NW India	na
P. scotica† Hooker	N Scotland	2n = 6x = 54
P. specuicola† Rydb.	SE Utah, N Arizona	2n = 2x = 18
P. stricta† Hornem.	Scandinavia, eastward to Kola Penins. (Russia), Novaya Zemlya (Russia), N Iceland, W Greenland, NE Canada	2n = 10-16x = ca. 88-136
P. yuparensis Takeda	Hokkaido (Japan)	2n = 4x = 36

†Species included in the present study. Source: Bruun (1932), Hultgård (1990; 1993), Kelso (1991), Richards (2002), Vogelmann (1956), www-sbras.nsc.ru/win/elbib/atlas/flora/4129.html

The *Aleuritia* complex comprises at least five different ploidy levels, including the diploids *P. alcalina*, *P. anvilensis*, *P. exigua, P. farinosa*, *P. frondosa*, *P. mistassinica*, *P. modesta* and *P. specuicola* (2n = 18), the tetraploids *P. borealis*, *P. halleri*, *P. sachalinensis* and *P. yuparensis* (2n = 36), the hexaploids *P. incana* and *P. scotica* (2n = 54), the octoploids *P. laurentiana*, *P. magellanica* and *P. scandinavica* (2n = 72), and the 14-ploid *P. stricta* (2n = ca. 126; Table 1; Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991; Hultgård, 1993; Richards, 2002). Chromosome numbers are unavailable for three Asian endemics: *P. baldschuanica*, *P. capitellata* and *P. schlagintweitiana* (Table 1). It has been suggested that most European polyploids of *Aleuritia* derive from a widespread, diploid *P. farinosa*-like ancestor (Bruun, 1932; Hultgård, 1990,

1993), while most American polyploids, including the South-American octoploid *P. magellanica*, from a widely distributed, diploid *P. mistassinica*-like progenitor (Kelso, 1991, 1992).

Despite recent studies on *Primula* (Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Conti *et al.*, 2000; Mast *et al.*, 2001; Trift *et al.*, 2002; Zhang *et al.*, 2004; Zhang and Kadereit, 2004; Guggisberg *et al.*, 2006b; Mast *et al.*, 2006; Carlson *et al.*, 2008), no investigations have yet aimed at elucidating patterns of hybrid speciation in the genus. Here, we intend to reconstruct the reticulate history of *Aleuritia* by using a combination of phylogenetic and consensus network analyses of chloroplast DNA (cpDNA) and ITS data, examination of polymorphic sites in ITS sequences, and identification of recombinant ITS clones. Specifically, we attempt to: 1) identify the putative parental lineages of the polyploids and compare them with those previously proposed on the basis of chromosome numbers, morphology, and distribution; 2) discern whether the polyploids evolved through auto- or allopolyploidisation; 3) establish whether polyploid taxa formed recurrently; and 4) assess whether *Aleuritia* evolved within the framework of the secondary contact model. A more general goal of our study is to illustrate how the use of different analytical approaches is fundamental to discern among the different possible sources of phylogenetic incongruence and provide a detailed scenario for the evolution of complex polyploid groups.

Materials and Methods

Taxon sampling

A total of 63 plant accessions were used for the present study (Appendix 1). Forty-four accessions represented 15 of the 21 species ascribed to *Aleuritia* and 19 accessions represented four of the 14 species ascribed to sect. *Armerina* (hereafter called *Armerina*; Table 1; Richards, 2002). The *Armerina* accessions were included because cytological, morphological and biogeographic evidence suggested that *P. egaliksensis*, assigned to *Armerina* (Richards, 2002), might have a hybrid origin involving a member of *Aleuritia* (i.e. *P. mistassinica*) and a member of *Armerina* (i.e. *P. nutans*; Kelso, 1991). Furthermore, recent phylogenies of *Primula* based on cpDNA markers supported the inclusion of selected members of *Armerina* and *Aleuritia* in the same, largely unresolved clade and, more specifically, of *P. egaliksensis* within a well-supported *Aleuritia* sub-clade (Mast *et al.*, 2001; Guggisberg *et al.*, 2006b; Mast *et al.*, 2006). Finally, *P. ioessa* (sect. *Sikkimensis*) was chosen to root the resulting trees, because members of the mentioned section were sister to a large clade comprising *Aleuritia* and *Armerina* in the latest *Primula* phylogeny (Mast *et al.*, 2006).

DNA extraction, PCR amplification, sequencing and cloning

Total genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Switzerland). Polymerase chain reactions (PCR) were performed in 20 µl volumes containing 1X buffer

(including 1.5 mM $MgCl_2$), 2 mM $MgCl_2$, 200 µM dNTPs, 0.2 µM of each primer, and one unit Taq polymerase (Sigma, Switzerland). Amplifications were carried out on a thermocycler (Biometra, Germany), using the following conditions: a first cycle at 94°C for 2 min; 35 cycles at 94°C for 30 s, 52°C for 1 min and 72°C for 1.75 min; a final cycle of 10 min at 72°C, while lowering the ramp speed from 5.0°C/s to 1.0°C/s for the annealing step.

Data matrices were generated from both the chloroplast and the nuclear genomes, respectively. The maternal inheritance of the cpDNA has been demonstrated in many angiosperms, including *Primula* (Corriveau and Coleman, 1988), whereas nuclear markers are biparentally inherited (Baldwin *et al.*, 1995; Álvarez and Wendel, 2003; Small *et al.*, 2004). Five cpDNA markers were each targeted with independent PCR reactions. The *rpl*16 intron was amplified with primers rpL16F71 and rpL16R1516 (Small *et al.*, 1998); the *rps*16 intron with primers rpS16F and rpS16R2 (Oxelman *et al.*, 1997); the *trn*L-F region (comprising the *trn*L intron, the *trn*L exon, and the *trn*L-*trn*F intergenic spacer) with primers 5'trnL^{UAA}F (TabC) and trnF^{GAA} (TabF; Taberlet *et al.*, 1991); the *trn*T-*trn*L intergenic spacer with primers trnT^{UGU}F (TabA) and 5'trnL^{UAA}R (TabB; Taberlet *et al.*, 1991); and the *trn*D-T region (comprising three intergenic spacers and the *trn*Y and *trn*E exons located between the *trn*D and *trn*T genes) with primers trnD^{GUC}F and trnT^{GGU} (Demesure *et al.*, 1995). The ITS region (comprising the ITS1, the 5.8S gene, and the ITS2) of the nrDNA cistron was amplified with primers ITS.LEU and ITS4 (Baum *et al.*, 1998). Sequences of the *trn*D-T and ITS region, including 302 ITS clones (see below), were generated for this study, for a total of 428 new sequences. The length of the amplified fragments was estimated by comparison with DNA ladders on 1.2% agarose gels stained with ethidium bromide. The absence of contamination was checked by running out the PCR products of a negative control on the same gel. Successfully amplified PCR products (amplicons) were cleaned with the GFX PCR DNA and Gel Band Purification Kit (Biosciences Amersham, Switzerland).

Sequencing reactions were prepared with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA), using the same primers as in the PCR amplifications. Sequencing products were purified on 96-well multiscreen filtration plates (Millipore, USA) to remove excess dye terminators and run on an ABI Prism 3100 automated sequencer (Applied Biosystems, USA). Sequencher 4.2 (Gene Codes Corp., USA) was used to check the quality of the electropherograms and compile the contiguous sequences (contigs) for each amplicon.

Direct sequencing of ITS amplicons produced electropherograms with double peaks and non complementarity between sequenced strands in the following accessions of *Aleuritia* and *Armerina*: borealis2, egaliksensis6, farinosa1/3/5/6/8, fasciculata, halleri2, incana1/2/3, laurentiana4,

mistassinica3/4/6/8/9, nutans1/3/4/5/6/7/8/9, scandinavica2, scotica1, stricta1/2. To confirm the nucleotide heterogeneity detected via direct sequencing of ITS amplicons and search for rare repeats in apparently homogeneous samples, PCR products from the following 17 accessions were cloned into pCR®II-TOPO® (TOPO TA Cloning® kit, Invitrogen, Basel) according to the manufacturer's instructions: alcalina, borealis2, egaliksensis7, farinosa1, halleri2, incana1, laurentiana1/4, mistassinica9, magellanica2, nutans7, scandinavica1/2, scotica1/2, stricta1/3. Despite the presence of double peaks, the ITS amplicon of fasciculata was not cloned, for *Armerina* was not the focus of the present study. Because preferential amplification of one sequence variant may lead to its over-representation in the final reaction mixture (cf. PCR drift and PCR selection; Wagner *et al.*, 1994), three PCR reactions of each sample were pooled for use in ligations (Mason-Gamer, 2004). Ten to 24 white colonies per cloning reaction, resulting in a final number of 302 clones, were randomly selected and their ITS inserts amplified prior to sequencing.

To estimate the percent of recombinant clones that might be attributed to PCR artefacts, total DNAs from *P. farinosa* (2x) and *P. halleri* (4x) were mixed at equal ratio for consecutive PCR and cloning (hereafter called "mixed cloning experiment"; see also Kovarík *et al.*, 2005) and the resulting percent of recombinants was compared with the corresponding values obtained for *P. scotica* (6x), *P. scandinavica* (8x) and *P. stricta* (accessions 1/2; 14x), because the diploid (*P. farinosa*) and tetraploid (*P. halleri*) lineages were hypothesised to have contributed to the origin of the higher polyploids (see Results; Bruun, 1932; Hultgård, 1990, 1993; Arnold and Richards, 1998; Richards, 2002).

Phylogenetic analyses

The starting and ending points of the sequences generated from each PCR amplification were determined by comparison with the complete cpDNA sequence of *Nicotiana tabacum* (GenBank Z00044) and the partial nrDNA sequence of *Rhododendron kanehirai* (GenBank AF172290). Nucleotide sequences were aligned by eye and 54bp from the *trn*D-T region had to be eliminated due to ambiguity in the alignment. Unequivocal gaps in cpDNA alignments were coded as present or absent according to Simmons and Ochoterena (2000) with the software GapCoder (Young and Healy, 2003) and added as binary characters to the end of the four-state nucleotide matrix. All analyses of cpDNA sequences described below were performed on the gapped matrices. The aligned data matrices and resulting phylogenetic trees were submitted to TreeBASE (www.treebase.org).

The matrices generated from each of the five cpDNA markers were first analysed separately under maximum parsimony (MP) optimisation using the beta 10 version of PAUP* 4.0 (Swofford, 1999; Table 2). MP trees for each dataset were generated by performing heuristic searches with

character states weighted equally, gaps/polymorphisms treated as missing data/uncertainties, TBR branch swapping, Steepest Descent ON, Mulpars ON, and Collapse branches option ON for branches with a minimum length of zero. Two hundred searches were performed under these conditions, after randomising the order of taxon addition. Five trees per replicate were saved and used as starting trees for a further round of branch swapping with TBR, now saving all the trees. Bootstrap support (BS) values (Felsenstein, 1985) for individual branches of the resulting trees were obtained from heuristic searches of 1000 bootstrap replicates with 100 random sequence addition (holding ten trees at each step), TBR branch swapping, Steepest Descent ON, and maxtrees set to 10.

Table 2. Character metrics and tree statistics from parsimony analyses of the cpDNA and nrDNA matrices. Total, total length of matrix (number of gaps); Informative, number of informative characters in the aligned matrix; Uninformative, number of uninformative characters in the aligned matrix; PICs, number of potentially informative characters; % Variability, calculated by dividing the PICs value by the length of the aligned matrix; N, number of trees; L, tree length; CI, consistency index; RI, retention index; and RC, rescaled coefficiency index.

	cpDNA						nrDNA
	*rpl*16 intron	*rps*16 intron	*trn*L-F region	*trn*T-L spacer	*trn*D-T region	Global dataset	ITS region
Character metrics:							
Total (gaps)	1011 (7)	864 (4)	1060 (8)	888 (4)	965 (4)	4788(27)	638
Informative	33	33	32	38	30	167	62
Uninformative	29	23	26	31	23	132	49
PICs	62	56	58	69	53	299	111
% Variability	6.13	6.48	5.47	7.77	5.49	6.24	17.40
Tree statistics:							
N	1	1	4	1	3	12	349
L	69	61	62	75	60	331	145
CI	0.942	0.951	0.968	0.947	0.9	0.934	0.828
RI	0.986	0.989	0.994	0.989	0.977	0.985	0.967
RC	0.929	0.941	0.962	0.936	0.879	0.92	0.801

Visual comparisons of the MP trees derived independently from each of the five cpDNA markers revealed no strongly supported (BS > 68%) topological inconsistencies. Therefore, cpDNA data partitions were combined into a "global cpDNA matrix" that was analysed using the same settings described above (Table 2). Conversely, strongly supported (BS = 100%) topological inconsistencies were identified by comparing the global cpDNA tree (Fig. 1) with the tree derived from MP analyses of the nrDNA sequences (Fig. 2) performed as described above (cf. Table 2), thus precluding the merging of cpDNA and ITS datasets for phylogenetic analysis of the combined dataset.

Bayesian analyses of the cpDNA and nrDNA sequences were implemented in MrBayes version 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Prior to analysis, datasets were partitioned (Nylander *et al.*, 2004) and the Akaike information criterion was used to select the best-fit models of nucleotide substitutions for each defined partition with Modeltest

version 3.6 (Posada and Crandall, 1998; Table 3). For the five cpDNA markers, the *trn*L exon, separating the *trn*L intron from the *trn*L-*trn*F intergenic spacer, was excluded because it was invariable, and the *trn*L-F region was split into two partitions: the *trn*L intron and the *trn*L-*trn*F intergenic spacer. Conversely, the two short exons (*trn*Y: 84 bp; *trn*E: 73 bp) separating the three spacers of the *trn*D-T region were retained in the data matrix, because they varied at one nucleotide position. In conclusion, the cpDNA characters were divided into the following partitions: *rps*16 intron, *rpl*16 intron, *trn*L intron, *trn*L-*trn*F intergenic spacer, *trn*T-*trn*L intergenic spacer, *trn*D-T region, and coded gaps. A binary model (Lset coding = variable) was applied to all gaps coded in the five cpDNA markers. The ITS region dataset was divided into the following categories: ITS1, 5.8S gene, and ITS2. The 5.8S gene, separating the ITS, was retained in the data matrix, because it varied at two nucleotide positions.

Posterior probabilities (PP) of the trees and parameters for the selected substitution models were approximated by MCMC simulations using one cold and three incrementally heated chains (Temp = 0.2) over $5 \cdot 10^6$ generations (Nylander *et al.*, 2004). Trees were sampled every 100 generations, resulting in 50,000 trees, minus the trees sampled during the burn-in phase (i.e. before chains reached stationarity). Stationarity was determined by plotting the log likelihood (lnL) against the number of generations (Gen) and ascertained for each parameter using Tracer v1.4 (Rambaud and Drummond, 2007). Two independent runs (Nruns = 2), starting from different random trees, were performed to ensure that the individual runs had converged to the same result. Convergence of runs was asserted when the potential scale reduction factor provided by the sump output approached 1.0, and the resulting tree topologies, branch lengths, and clade credibility values were similar across runs (Huelsenbeck and Ronquist, 2001; Huelsenbeck *et al.*, 2002; Ronquist and Huelsenbeck, 2003). Final inference of the trees and parameters was obtained from the concatenation of the two runs (Table 4).

Visualisation of topological conflicts

To facilitate comparison of the phylogenetic conflicts supported by the cpDNA and nrDNA sequences, we mirrored the respective 50% majority-rule consensus trees estimated by MP (cf. Figs. 1-2) after editing them as follows (Fig. 3a): (i) for monophyletic species, multiple accessions were collapsed to a single terminal, while for non-monophyletic species, accessions representing different clades in either tree were maintained (cf. Figs. 1-2); (ii) nodes with BS lower than 70% were collapsed (cf. Figs. 1-2), ensuring the comparison of only well-supported topological inconsistencies. The resulting trees are called “simplified 70% BS consensus trees” from now on (Fig. 3a).

Table 3. Best-fitting substitution models and parameter values for the separate data partitions. Model abbreviations as in MrBayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The *trn*L exon was excluded, because it was invariable. A dash indicates that the parameter value was not estimated. I, proportion of invariable sites assumed by the model; Γ, gamma distribution assumed by the model to describe among sites rate-variation.

Data partition	**Substitution Model**										
	Model	Base freq.				Subst. rates (relative to G<->T rate)					I
		A	C	G	T	A<->C	A<->G	A<->T	C<->G	C<->T	
*rpl*16 intron	K81uf+I	0.4121	0.1416	0.1583	0.288	1	0.9224	0.1082	0.6753	0.9224	0.6753
*rps*16 intron	TVM+Γ	0.3646	0.1448	0.1772	0.3134	1.143	1.3832	0.0444	0.5612	1.3832	-
*trn*L intron	K81uf+I	0.3883	0.1331	0.158	0.3206	1	0.6301	0.1912	0.8144	0.6301	0.8144
*trn*L-*trn*F interg. spacer	TVM	0.2914	0.172	0.1411	0.3955	0.3984	0.8765	0	0.2717	0.8765	-
*trn*T-*trn*L interg. spacer	TVM	0.393	0.1207	0.1369	0.3494	0.9434	0.5603	0.0144	0.2656	0.5603	-
*trn*D-*trn*T region	K81uf+I	0.3236	0.1708	0.1555	0.3501	1	0.6176	0.1156	0.7784	0.6176	0.7784
ITS1	TIMef+Γ	equal				1	1.9478	0.4263	0.4263	4.6371	-
5.8S gene	JC	equal						equal			-
ITS2	TrN+I	0.19	0.2631	0.2605	0.2864	1	3.0662	1	1	5.6637	0.319

Table 3 (continued)

Data partition	**Bayesian parameters**			
	Γ	Prior state freq.	Nb. subst.	Among site var.
*rpl*16 intron	-	estimated	6	propinv
*rps*16 intron	0.3681	estimated	6	gamma
*trn*L intron	-	estimated	6	propinv
*trn*L-*trn*F interg. spacer	-	estimated	6	equal
*trn*T-*trn*L interg. spacer	-	estimated	6	equal
*trn*D-*trn*T region	-	estimated	6	propinv
ITS1	0.6556	estimated	6	gamma
5.8S gene	-	fixed (equal)	1	equal
ITS2	-	estimated	6	propinv

Table 4. Bayesian posterior parameters and tree statistics for the cpDNA and nrDNA datasets. TL, total tree length; I, proportion of invariable sites assumed by the model; Γ, gamma distribution assumed by the model to describe among sites rate-variation.

Parameters		**cpDNA dataset**	**nrDNA dataset**
Nb. samples (2 runs)		98802	98802
TL		0.105495	0.466732
Subst. model		GTR+I+Γ	GTR+I+Γ
Base freq.	A	0.366398	0.219585
	C	0.147621	0.269826
	G	0.158147	0.260432
	T	0.327834	0.250157
Subst. rates (relative to G<->T rate)	A<->C	0.966467	1.395558
	A<->G	0.740649	2.647155
	A<->T	0.054081	1.722177
	C<->G	0.342878	0.431228
	C<->T	0.855861	6.143414
I		0.766699	0.482981
Γ		87.056056	7.091401

Additionally, consensus networks were computed, because they allow for the simultaneous visualisation of competing hypotheses of relationships supported by a group of trees (Holland and Moulton, 2003; Holland *et al.*, 2005). Consensus networks may vary between tree-like representations, when trees are fully congruent with each other, and net-like representations, when trees are strongly incongruent with each other (Holland and Moulton, 2003; Holland *et al.*, 2005). The complexity of a network also depends on the threshold *x*, representing the branches (or splits σ) that occur in a certain proportion of the trees. For example, a network calculated by using $x = 0.1$ displays the splits occurring in at least 10% of the trees. Thus, lower or higher values of *x*, respectively, capture a greater or smaller proportion of the topological conflicts in a group of trees (McBreen and Lockhart, 2006).

To investigate the topological inconsistencies among the trees generated from the global cpDNA dataset, consensus networks were estimated from the 12 MP and the almost 100,000 BI trees (Tables 2 and 4), respectively, with $x = 0.1$ in SplitsTree 4.6 (Fig. 4a; Huson, 1998; Huson and Bryant, 2006). Because the high number of trees sampled by BI led to software crash, it was reduced by eliminating duplicate topologies (CONDENSE command) and collapsing branches with a minimum length of zero (Collapse = MinBrLen) in PAUP*4.0b10 (Swofford, 1999). Similarly, consensus networks were estimated from the 349 MP and the almost 100,000 BI trees, respectively, (Tables 2 and 4) generated from the nrDNA dataset (Fig. 4b).

To identify the topological inconsistencies caused by diploid and polyploid accessions, respectively, consensus networks were computed from the simplified 70% BS consensus trees of the cpDNA and nrDNA matrices (cf. Fig. 3a) by either excluding (using the Data→Filter Taxa option in SplitsTree 4.6; Fig. 3d; Huson, 1998; Huson and Bryant, 2006) or including the polyploids in the analyses (Fig. 3e). We kept the same value of *x* (0.1) as above, because, when comparing only two trees, any *x* value between 0 and 0.5 will ensure representation of all topological conflicts.

Identification of additive polymorphic sites

We defined any position of the ITS alignment that varied among direct sequences from amplicons of different accessions as a "variable" site (e.g. site 21 in Appendix 2). A site was designated as "polymorphic" when more than one peak was present in the electropherogram derived from the direct sequence of an ITS amplicon (Appendix 2), or when at least two different nucleotides were detected at the same position among the clones of a single ITS amplicon (Appendix 3). In either case, the different nucleotides detected at the same position in the same individual are defined as polymorphic nucleotides, forming a single polymorphism (i.e. infra-individual variation). Given a rate of $2 \cdot 10^{-4}$ errors/bp and an average length of 638bp for each sequence, the 302 sequenced ITS clones may contain 38 point mutations resulting from Taq error. Hence, nucleotide positions

differing in only one clone per amplicon were not defined as polymorphic. For example, while the direct sequence of an amplicon from sample scotica1 presented six polymorphic sites (e.g. position 11 of the ITS alignment; Appendix 2), cloning of an amplicon from the same accession revealed 23 polymorphic sites, 17 of which were not detected by direct sequencing (e.g. position 12 of the ITS alignment; Appendix 3). In addition, 18 positions that differed in only one clone were identified and, therefore, not defined as polymorphic (e.g. position 45, clone5; Appendix 3).

Polymorphisms are commonly treated as uncertainties in phylogenetic analyses, but, when they consist of nucleotides also found in other accessions of the dataset (i.e. additive nucleotides), they may help to identify the progenitors of hybrid species (Sang *et al.*, 1995; Campbell *et al.*, 1997). The polymorphic sites containing additive nucleotides are defined as additive polymorphic sites (APS; Fuertes Aguilar and Nieto Feliner, 2003). Hence, to disentangle putative reticulation patterns in *Aleuritia*, the matrices of polymorphic sites derived from both direct and cloned sequences (online Appendices 2-3) were searched by eye and all the APS's of the ITS alignment compiled in Table 5.

Analyses of variation among ITS clones

To investigate the infra-individual variation among cloned ITS sequences and visualise the positions of recombinant sequences of the higher polyploids in relation to lower polyploids and diploids, multivariate analyses were performed using NTSYSpc version 2.11S (Rohlf, 2000). Pairwise comparisons calculated with PAUP* 4.0b10 (Swofford, 1999) for all possible pairs of clones by means of absolute distances (i.e. Manhattan metric) between sequences generated triangular similarity matrices that were used for principal coordinates analysis (PCoA; Gower, 1966).

Initially, all 179 sequenced ITS clones (excluding 123 duplicates) were analysed by PCoA. Because this first analysis revealed that clones from the European clade A (cf. Fig. 2) did not form a tight cluster, the respective subset of 91 clones (excluding duplicates) was analysed separately to increase resolution (Fig. 5a). Since this second PCoA showed that ITS sequences from the high polyploids of clade A (i.e. *P. scotica* [6x], *P. scandinavica* [8x], and sample stricta1 of *P. stricta* [14x]) occupied intermediate positions between those of *P. farinosa* (2x) and *P. halleri* (4x; Fig. 5a), we further investigated the origin of the mentioned polyploids by mapping the base composition at the ten polymorphic sites that differ between *P. farinosa* and *P. halleri* (i.e. positions 32, 43, 61, 85, 105, 167, 208, 231, 235 and 514 of the ITS alignment ; see Table 5) onto diagrams of the ITS sequences of *P. scotica* , *P. scandinavica* , and sample stricta1 of *P. stricta* (Fig. 5b). A recombination event was inferred each time there was a switch (identified by a cross on Figure 5b)

from a nucleotide specific to *P. farinosa* to another specific to *P. halleri*, or vice-versa (see Appendix 3; Buckler *et al.*, 1997).

Table 5. Thirty-two additive polymorphic sites (APS) from the ITS region of *Primula* sect. *Aleuritia*. Additive nucleotides are highlighted in grey. The ten APS at which *P. farinosa* and *P. halleri* differ are designated by arrows. Site 378 belonging to the 5.8S gene and separating the ITS1 from ITS2 is italicised. Clade affiliation as in Figure 2.

				↓	↓		↓	↓	↓							↓		
									1	1	1	1	1	1	1	1	1	
		1	1	3	4	5	6	8	0	0	2	2	5	6	6	7	9	
Clade	**Accessions**	1	2	2	3	4	1	5	5	9	3	5	6	0	7	5	3	
'Asian-American'	*P. mistassinica*	T	G	C	A	A	T	G	C	C	G	G	C	C	G	C	G	
	P. specuicola	T	A	C	A	A	C	G	C	C	G	G	C	C	G	C	G	
	P. alcalina	T	A	C	A	A	C	G	C	C	T	G	C	C	G	C	G	
	P. anvilensis	C	A	C	A	A	C	G	C	C	G	G	C	C	G	C	G	
	P. modesta	T	A	C	A	A	C	G	C	C	C	G	C	C	G	C	G	
	P. borealis (4x)	A	A	C	A	A	C	G	C	C	G	G	C	C	G	C	G	
	P. incana (6x)	A	A	C	A	A	C	G	C	C/T	C/G	G	C	C	G	C/T	G	
	P. laurentiana (8x)	A	A	C	A	A	C	G	C	C	G	G	C	C	G	C	G	
	P. magellanica (8x)	T	A	C	A	A/G	C	G	C	C	G	G	C	C	G	C	G	
	*P. stricta*3 (14x)	A	A	C	A	A	C	G	C	C	G	G	C	C	G	C	G	
'European'	*P. farinosa*	T	A	T	A	A	T	G	C	A/T	G	G/−	C	C	A	T	A/T	
	P. halleri (4x)	T	A	C	C	A	C	C	T	T	G	G	C/T	C	G	T	A	
	P. scotica (6x)	A/G/T	A/C	C/T	A/C	A	C/T	C/G	C/T	T	G	A/G	C	A/C	A/G	T	A	
	P. scandinavica (8x)	T	C/G	C/T	A/C	G	C/T	A/G	C/T	T	G	G	C	C	A/G	T	A	
	*P. stricta*1/2 (14x)	A/T	A/G	C/T	A/C	A	C/T	A/G	C/T	T	G	G/−	C/T	A/C	A/G	T	A	

Table 5 (continued)

					↓		↓	↓						↓				
		1	1	2	2	2	2	2	*3*	4	4	5	5	5	5	5	6	6
		9	9	0	0	1	3	3	*7*	4	4	0	0	1	7	9	0	1
Clade	**Accessions**	3	8	2	8	9	1	5	*8*	5	6	0	4	4	4	4	6	7
'Asian-American'	*P. mistassinica*	G	A	C/T	T	C	G	C	*T*	C	C	C	C	A	T	A/G	A	T
	P. specuicola	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C	G	A	C
	P. alcalina	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C	G	A	C
	P. anvilensis	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C	G	A	C
	P. modesta	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C	G	A	C
	P. borealis (4x)	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C/T	G	A	C/T
	P. incana (6x)	G	A	C	T	C	G	C	*C*	C	C	C	C	A	T	G	A	C
	P. laurentiana (8x)	G	A	C	T	C	G	C/G	*C*	C	C	C	C	A	T	G	A	C
	P. magellanica (8x)	G	A	C	T	C	G	C	*C*	C	C	C	C	A	C	G	A	C
	*P. stricta*3 (14x)	G	A	C	T	C	G	C	*C*	C	C	C	C	A	T	G	A/G	C
'European'	*P. farinosa*	A/T	A	C	T	C	G	G	*C*	C	C	C	C	A	T	G	A/G	C
	P. halleri (4x)	A	A	C/T	A	C	T	C	*C*	C	C	C	C	T	T	A/G	A	C
	P. scotica (6x)	A	A/C	C	A/T	C/T	G/T	C/G	*C/T*	A/C	C/T	C	C/T	A/T	T	A/G	A	C
	P. scandinavica (8x)	A	A	C/T	A/T	C	G	C/G	*C/T*	A/C	C	C/T	C	A	T	A/G	A	C
	*P. stricta*1/2 (14x)	A	A/C	C/T	A/T	C/T	G	C/G	*C*	C	C/T	C/T	C/T	A	T	G	A	C

Results

Intersectional hybridisation

The consensus networks computed from the simplified 70% BS consensus trees yield little incongruence inherent to diploid species (Fig. 3d). Most of the topological conflicts between the nrDNA and cpDNA trees are caused by polyploid lineages (Fig. 3e), in particular the tetraploid, North-American *P. egaliksensis* (2n=40; Fig. 3a, b) which forms a strongly supported clade with the diploid, North-American *P. mistassinica* (2n=18; Fig. 3b) of *Aleuritia* in the cpDNA tree (clade V, Fig. 1) and the diploid, circumboreal *P. nutans* (2n=22) ascribed to *Armerina* in the nrDNA tree (Fig. 2). This evidence and the observation that the chloroplast genome is maternally inherited in Primulaceae (Corriveau and Coleman, 1988) suggest that *P. egaliksensis* is an intersectional hybrid resulting from allopolyploidisation between a *P. mistassinica*-like and a *P. nutans*-like ancestor, the former providing the maternal parent (Fig. 6). Furthermore, multivariate analyses of cloned PCR products showed that all ITS sequences from *P. egaliksensis* clustered with *P. nutans* (data not shown), confirming that they had homogenised towards a single repeat-type.

Origins of polyploids belonging to the Asian-American clade

In the nrDNA tree, all accessions of the North-American polyploids *P. incana* (6x) and *P. laurentiana* (8x) form a weakly supported polytomy with the amphi-Beringian *P. borealis* (4x) and accession stricta3 of the amphi-Atlantic *P. stricta* (14x; clade B, Fig. 2), while the mentioned taxa are dispersed over three clades (I, II, V) in the cpDNA tree (Fig. 1). The incongruent placement of these polyploids in the simplified 70% BS consensus trees (Fig. 3a, e) may imply a hybrid origin of these taxa.

The sister relationship of the amphi-Beringian tetraploid *P. borealis* (2n=36; Fig. 3b) with the diploid *P. anvilensis* (2n=18) endemic to the Alaskan Seward Peninsula (Fig. 3b) in the cpDNA tree, with the diploid *P. modesta* (2n=18) endemic to East Asia (Fig. 3b) as a close relative (clade I, Fig. 1), and its inclusion in a clade with *P. anvilensis* and *P. modesta* in the nrDNA tree (clade B, Fig. 2) suggest an allopolyploid origin from the mentioned diploids, with the *P. anvilensis* lineage possibly providing the maternal ancestor. However, *P. borealis* does not cause any incongruence in the consensus network computed from the simplified 70% BS consensus trees (Fig. 3e), reflecting the fact that its differing relationships in the cpDNA and nrDNA trees are not strongly supported (BS = 58% for clade B in the nrDNA tree; Fig. 2). Our results thus do not allow us to discern between a possible hybrid origin of *P. borealis* from *P. anvilensis*-like and *P. modesta*-like parents and an autopolyploid origin from a *P. anvilensis*-like ancestor (Fig. 6).

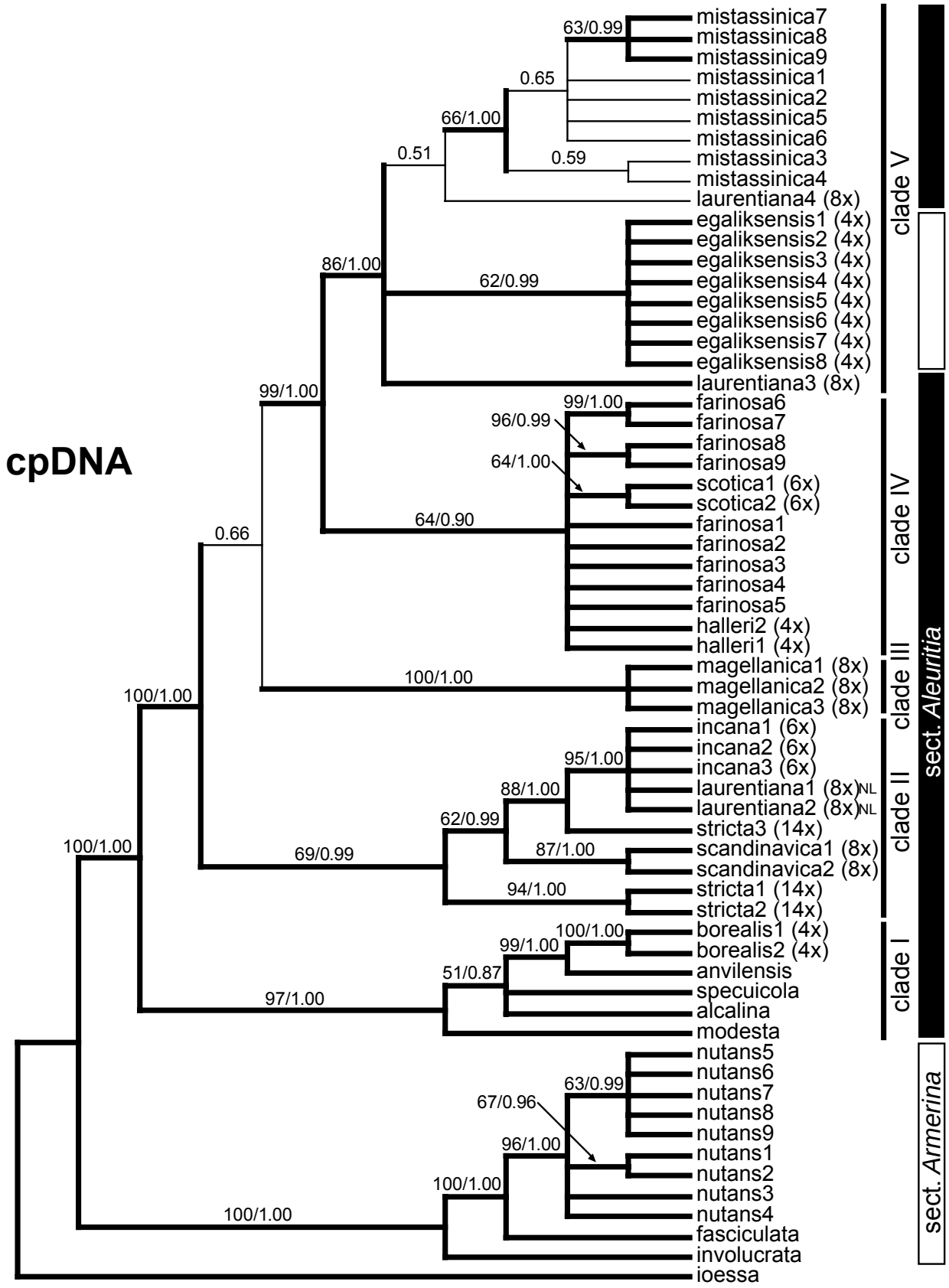


Figure 1. Fifty percent majority-rule consensus cladogram inferred by Bayesian analysis of the global cpDNA dataset. Clades recovered also in the 50% majority-rule consensus of the 12 MP trees are highlighted by thicker branches. Bootstrap support values greater than 50% and posterior probabilities, separated by a slash, are indicated above the branches of the cladogram. Major clades are named using Roman numerals; abbreviations for terminals are given in Appendix 1. NL, Newfoundland.

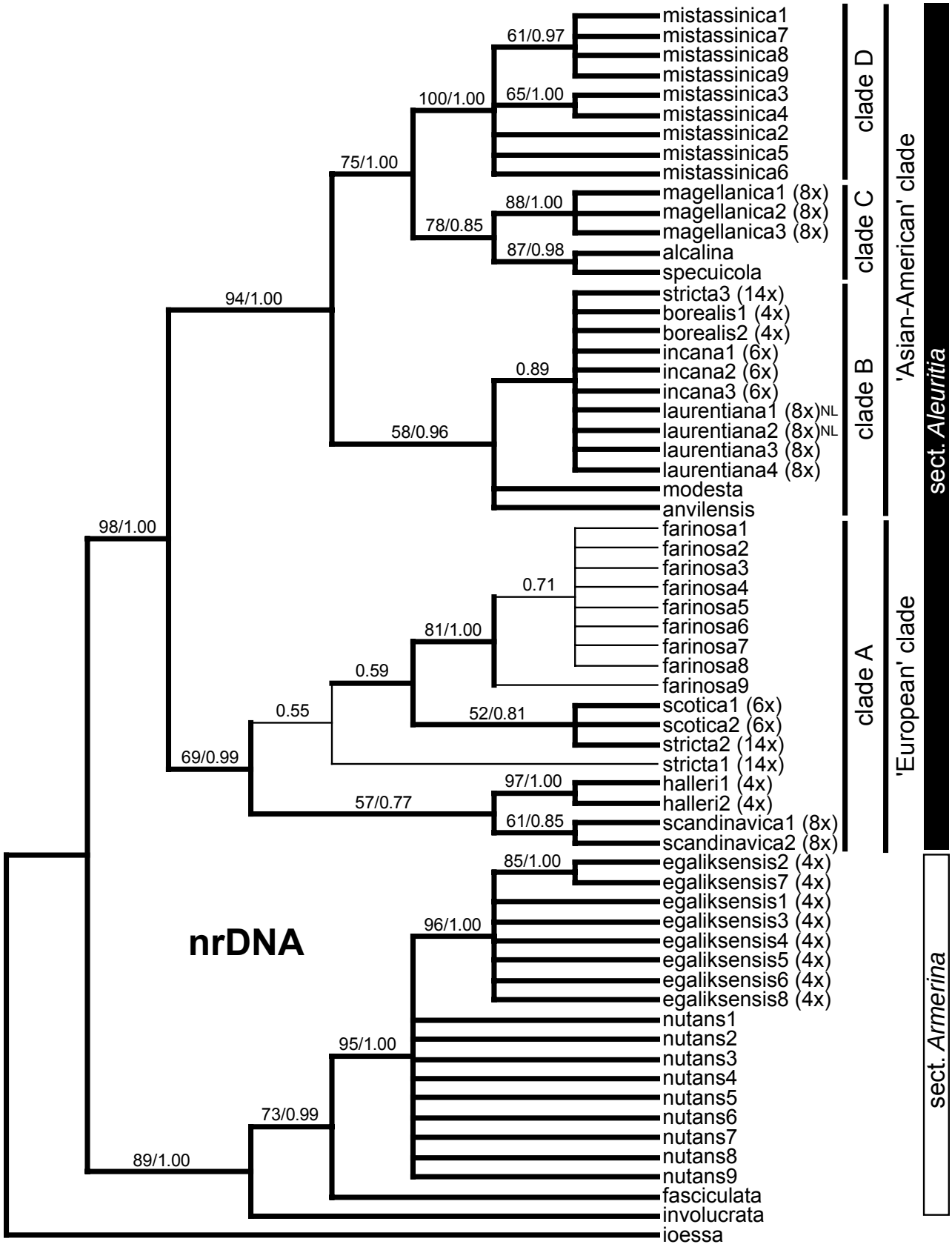


Figure 2. Fifty percent majority-rule consensus cladogram inferred by Bayesian analysis of the nrDNA dataset. Clades recovered also in the 50% majority-rule consensus of the 349 MP trees are highlighted by thicker branches. Bootstrap support values greater than 50% and posterior probabilities, separated by a slash, are indicated above the branches of the cladogram. Major clades are named using capital letters; abbreviations for terminals are given in Appendix 1. NL, Newfoundland.

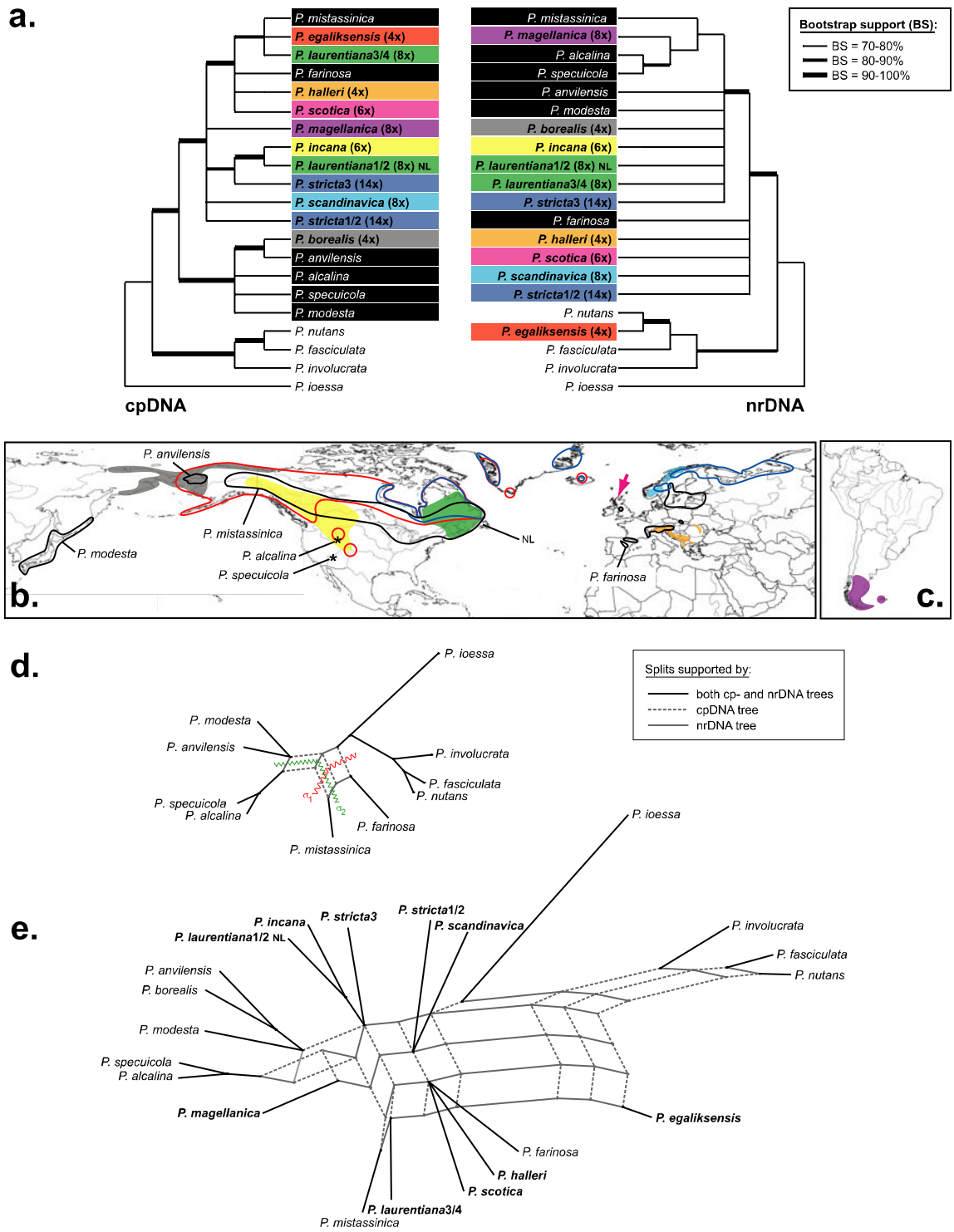


Figure 3. (a) Simplified 70% BS consensus trees supported by cpDNA (left) and nrDNA (right) datasets. (b, c) Distributional ranges of the tree terminals represented with matching colours. (d, e) Consensus networks computed from the two simplified cpDNA and nrDNA 70% BS consensus trees to identify topological inconsistencies caused by diploid (d) and polyploid accessions (e), respectively. Splits supported by both cpDNA and nrDNA trees, only the cpDNA tree or only the nrDNA tree, respectively, are indicated by plain black, grey dashed and plain grey lines. (d) By cutting through the splits supported by the cpDNA tree (e.g. σ_1: red wavy line) and the nrDNA tree (e.g. σ_2: green zigzag line), respectively, it is apparent that incongruence among the diploid species arises from the placement of *Primula mistassinica* with either *P. farinosa* in the former or with (*P. alcalina, P. specuicola*) in the latter (cf. Fig. 3a). (e) The topological conflicts caused by polyploid taxa mainly arise from the differential placement of *P. egaliksensis* and accessions of *P. incana*, *P. laurentiana*, *P. magellanica*, *P. scandinavica* and *P. stricta* between the nrDNA and cpDNA trees (cf. Fig. 3a). Polyploid accessions are in bold; abbreviations as in Appendix 1. NL, Newfoundland.

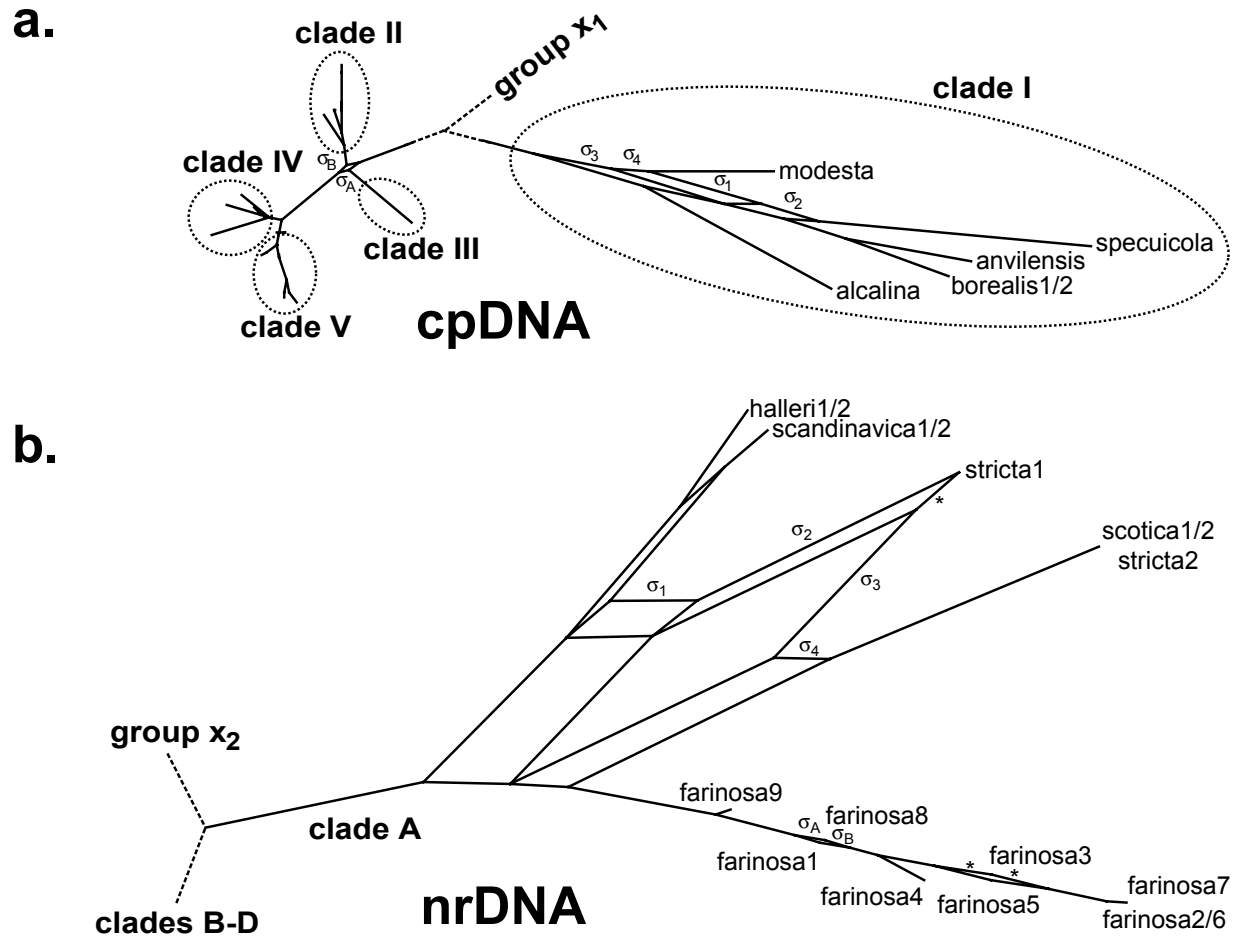


Figure 4. Consensus networks computed from the 12 and 349 MP trees, respectively, inferred from the cpDNA (a) and nrDNA (b) datasets. Splits not recovered in the consensus networks calculated from the corresponding Bayesian trees are indicated by an asterisk. The following zones of conflict were detected in both MP and BI networks: (a) splits σ_1-σ_4 and σ_A-σ_B, respectively, reflect the polytomies in clade I and between clades III and IV-V of the consensus tree of Figure 1; (b) splits σ_1- σ_4 and σ_A- σ_B, respectively, reflect the unresolved placement of stricta1 and farinosa9 in the consensus tree of Figure 2. Group x_1 represents accessions of *Primula nutans, P. fasciculata, P. involucrata* and *P. ioessa*. Group x_2 represents accessions of *P. egaliksensis*, *P. nutans, P. fasciculata, P. involucrata* and *P. ioessa*. Exceedingly long splits are represented by dotted lines. Clade names and abbreviations as in Figures 1-2 and Appendix 1.

The inclusion of the amphi-Beringian *P. borealis* (4x) in a polytomy with *P. incana* (6x), *P. laurentiana* (8x), and accession stricta3 of *P. stricta* on the nrDNA tree (14x; clade B, Fig. 2) and the tight clustering of cloned ITS sequences from these four taxa in the PCoA scatterplots (data not shown) suggest that the tetraploid likely played a central role in the origin of the higher polyploids. However, our current data do not allow us to speculate on the additional parent that might have contributed to the origin of the hexaploid *P. incana* (2n=54; Fig. 6) endemic to north-western North-America (Fig. 3b).

The results presented here, though far from being fully conclusive, provide useful evidence on the origin of *P. laurentiana*, an octoploid endemic to north-eastern North-America (2n=72; Fig. 3b). The four accessions of this species, while being included in a polytomy with the other North-American polyploids in the nrDNA tree (incl. the tetraploid *P. borealis* [2n=36] and the hexaploid

P. incana [2n=54]; clade B, Fig. 2), are dispersed over two distantly related clades in the cpDNA tree (Fig. 1): accessions laurentiana1/2, from Newfoundland (designated by NL on Figures 1-3), cluster with the hexaploid *P. incana* (2n=54; clade II), whereas accessions laurentiana3/4, from the Canadian mainland, are included in a clade with the diploid *P. mistassinica* (2n=18) and tetraploid *P. egaliksensis* (2n=40; clade V). The polyphyletic placement of *P. laurentiana* (8x) accessions in the cpDNA tree suggests that it might have originated independently in different locations. The Newfoundland accessions may result from hybridisation between a reduced gamete (2x) of a *P. borealis*-like ancestor and an unreduced gamete (6x) of a *P. incana*-like ancestor, the latter representing the maternal lineage (Figs. 1, 6). Conversely, Canadian mainland accessions may result from hybridisation between unreduced gametes of a *P. incana*- (6x) and a *P. mistassinica*-like (2x) lineage or via hybridisation between a reduced gamete (2x) of a *P. borealis*-like ancestor and an unreduced gamete (2x) of a *P. mistassinica*-like ancestor, followed by autopolyploidisation, the latter providing the maternal progenitor (Figs. 1, 6). The tetraploid *P. egaliksensis* (2n=40) cannot be excluded as a potential ancestor of *P. laurentiana* (2n=54; see clade V, Fig. 1), but its involvement is unlikely, because no species presenting 24 chromosomes can be invoked as the second progenitor, to fulfil the additivity in chromosome number.

The results of our phylogenetic analyses apparently support polyphyletic origins also for the amphi-Atlantic 14-ploid *P. stricta* (2n=126), which ranges from the eastern Canadian high Arctic, across Greenland and Russia, all the way to the north-eastern island of Novaya Zemlya (Fig. 3b), for different accessions (stricta1, 2 and 3) do not form a clade either in the cpDNA nor in the nrDNA tree (Figs. 1-2). The relationships of the North-American accession stricta3 in the cpDNA and nrDNA phylogenies do not contradict each other (Fig. 3e), for the accession forms a well-supported clade with the hexaploid *P. incana* (2n=54) and octoploid *P. laurentiana* (2n=72) in the cpDNA phylogeny (clade II, Fig. 1) and is included in the same polytomy with both species in the nrDNA tree (clade B, Fig. 2). Congruently, cloned ITS sequences of accession stricta3, *P. incana* and *P. laurentiana* cluster very tightly in the PCoA scatterplot (data not shown). These results therefore suggest that the North-American *P. stricta* derives from a *P. incana*- and a *P. laurentiana*-like lineage (Fig. 6).

Despite persistent uncertainties, the present study offers new insights into the origin of the octoploid *P. magellanica* (2n=72), the only species of *Primula* that occurs in the Southern Hemisphere, where it is restricted to Patagonia, Tierra del Fuego and the Falkland Islands (Fig. 3c). *Primula magellanica* groups with either clade II or clade IV-V in the cpDNA tree (Fig. 4a), and with the North-American diploids *P. alcalina*, restricted to north-eastern Idaho (2n=18; Fig. 3b), and *P. specuicola*, endemic to south-western United-States (2n=18; Fig. 3b), in the nrDNA tree

(clade C, Fig. 2), suggesting that the octoploid might derive from a *P. alcalina*- and/or *P. specuicola*- like ancestor (Fig. 6).

Origins of polyploids belonging to the European clade

All accessions of the European polyploids *P. halleri* (4x), *P. scotica* (6x) and *P. scandinavica* (8x) form a clade with accessions stricta1/2 of the amphi-Atlantic *P. stricta* (14x) and the European diploid *P. farinosa* in the nrDNA tree (clade A, Fig. 2), while the latter groups only with *P. halleri* and *P. scotica* in clade IV of the cpDNA phylogeny (Fig. 1), for *P. scandinavica* and stricta1/2 are included in clade II (Fig. 1). As already suggested for the North-American polyploids, the phylogenetic incongruence resulting from the variable position of the European polyploids on the simplified 70% BS consensus trees (Figs. 3a, e) advocates for a hybrid origin of these taxa. This assumption is corroborated by the observation that 60 out of 69 (86.96%) ITS additive nucleotides were found in the European clade (clade A, Fig. 2), namely four in the diploid *P. farinosa* and 56 in the polyploid taxa (Table 5).

The tetraploid *P. halleri*, endemic to the eastern Alps (2n=36; Fig. 3b), forms a poorly supported polytomy with the alpine diploid *P. farinosa* (Fig. 3b) in the cpDNA tree (2n=18; clade IV, Fig. 1) and is a member of a weakly resolved clade that also includes *P. farinosa* in the nrDNA tree (clade A, Fig. 2), suggesting that the widespread diploid contributed to the origin of the endemic tetraploid. However, our current data do not allow us to discern whether *P. halleri* originated via autopolyploidisation from a *P. farinosa*-like ancestor, or via allopolyploidisation involving an additional lineage (Fig. 6).

The hexaploid *P. scotica*, endemic to northern Scotland and adjacent Orkney Islands (2n=54; Fig. 3b), falls in the same clade as the diploid *P. farinosa* (2n=18) and the tetraploid *P. halleri* (2n=36) in both the cpDNA (clade IV, Fig. 1) and nrDNA (clade A, Fig. 2) trees, suggesting that the diploid and the tetraploid contributed to the origin of the hexaploid. This hypothesis is corroborated by ten polymorphic nucleotides that are additive for those distinguishing *P. farinosa* from *P. halleri*, respectively (Table 5). Furthermore, the analysis of cloned ITS sequences shows evidence of recombination between the ten sites differing between *P. farinosa* and *P. halleri* in 22 clones of *P. scotica* (excluding duplicates; Fig. 5b) which form a separate cluster close to the *P. farinosa*-clones (Fig. 5a), while two clones each were fully homogenised either towards the *P. farinosa* or the *P. halleri* repeat (Fig. 5b), the latter clustering with the respective groups in the PCoA scatterplot (i.e. scotica1 clone2/16, and scotica1 clone5/12, respectively; Fig. 5a). The combined evidence thus suggests that *P. scotica* derived from allopolyploidisation between a *P. farinosa*- and a *P. halleri*-like ancestor (Fig. 6).

The incongruent position of the octoploid *P. scandinavica* (2n=72) in the simplified 70% BS consensus trees (Fig. 3a, e) advocates for an allopolyploid origin of this endemic species to Norway (Fig. 3b), but phylogenetic evidence alone, while weakly supporting a sister relationship with the tetraploid *P. halleri* (2n=36) in the nrDNA tree (BS = 57%; clade A, Fig. 2), remains inconclusive regarding the specific parental ancestors of the octoploid. Detailed analyses of ITS sequences proved more informative. For example, *P. scandinavica* shares ten polymorphic nucleotides with the hexaploid *P. scotica* (2n=54), two of which (cf. positions 378 and 445) are made of nucleotides exclusive to these two species and eight can be traced back to the diploid *P. farinosa* (2n=18) and the tetraploid *P. halleri* (2n=36; cf. positions 32, 43, 61, 105, 167, 208, 235 and 594; Table 5). The analysis of 27 cloned ITS sequences (excluding duplicates) from two accessions of *P. scandinavica* reveals different degrees of recombination between sites differing between *P. farinosa* and *P. halleri* (Fig. 5b). The extensive recombination and the almost equal proportion of sites shared with *P. farinosa* and *P. halleri* in four (i.e. scandinavica1 clone6, scandinavica2 clone3/10/12; R! = 3-4, Fig. 5b) of the *P. scandinavica* ITS clones are reflected in the intermediate position they occupy in the PCoA scatterplot (Fig. 5a), whereas partial or complete homogenisation towards either the *P. farinosa* or *P. halleri* repeat types (Fig. 5b) links the *P. scandinavica* sequences with the respective group in the scatterplot (Fig. 5a). Combined evidence thus indicates that *P. scandinavica* (8x) might have originated either (i) via allopolyploidisation from unreduced gametes of a *P. scotica*- (6x) and a *P. farinosa*-like (2x) ancestor, or (ii) from an unreduced gamete (6x) of a *P. scotica*-like ancestor and a reduced gamete (2x) of a *P. halleri*-like ancestor, or (iii) via hybridisation between an unreduced gamete (2x) of a *P. farinosa*-like ancestor and a reduced gamete (2x) of a *P. halleri*-like ancestor, followed by autopolyploidisation (Fig. 6). However, the placement of *P. scandinavica* on the cpDNA tree (clade II, Fig. 1), far apart from remaining European species of lower ploidy level, may also indicate that the maternal parent has not been sampled (Hoot *et al.*, 2004).

Combined evidence from different analyses seems to corroborate an allopolyploid origin also for the European accessions of *P. stricta*. First, accessions 1 and 2 of the amphi-Atlantic 14-ploid *P. stricta* (2n=126) fall in the same clade as the octoploid *P. scandinavica* (2n=72) both in the cpDNA (clade II, Fig. 1) and in the nrDNA (clade A, Fig. 2) trees. Second, accession stricta2 forms a clade with the hexaploid *P. scotica* (2n=54) in the nrDNA tree, while accession stricta1 groups either with the clade containing *P. scotica* or with that of *P. scandinavica* (clade A, Fig. 4b). Third, accessions stricta1 and 2 share seven polymorphic nucleotides with *P. scotica* and *P. scandinavica* (Table 5). Finally, all except one (excluding duplicates; cf. stricta1 clone5/11, Fig. 5b) cloned ITS sequences of accession stricta1 are recombinant for sites differing between *P. farinosa* and *P. halleri* (Fig. 5b), and therefore cluster with either *P. scotica* or *P. scandinavica* sequences in the PCoA scatterplot

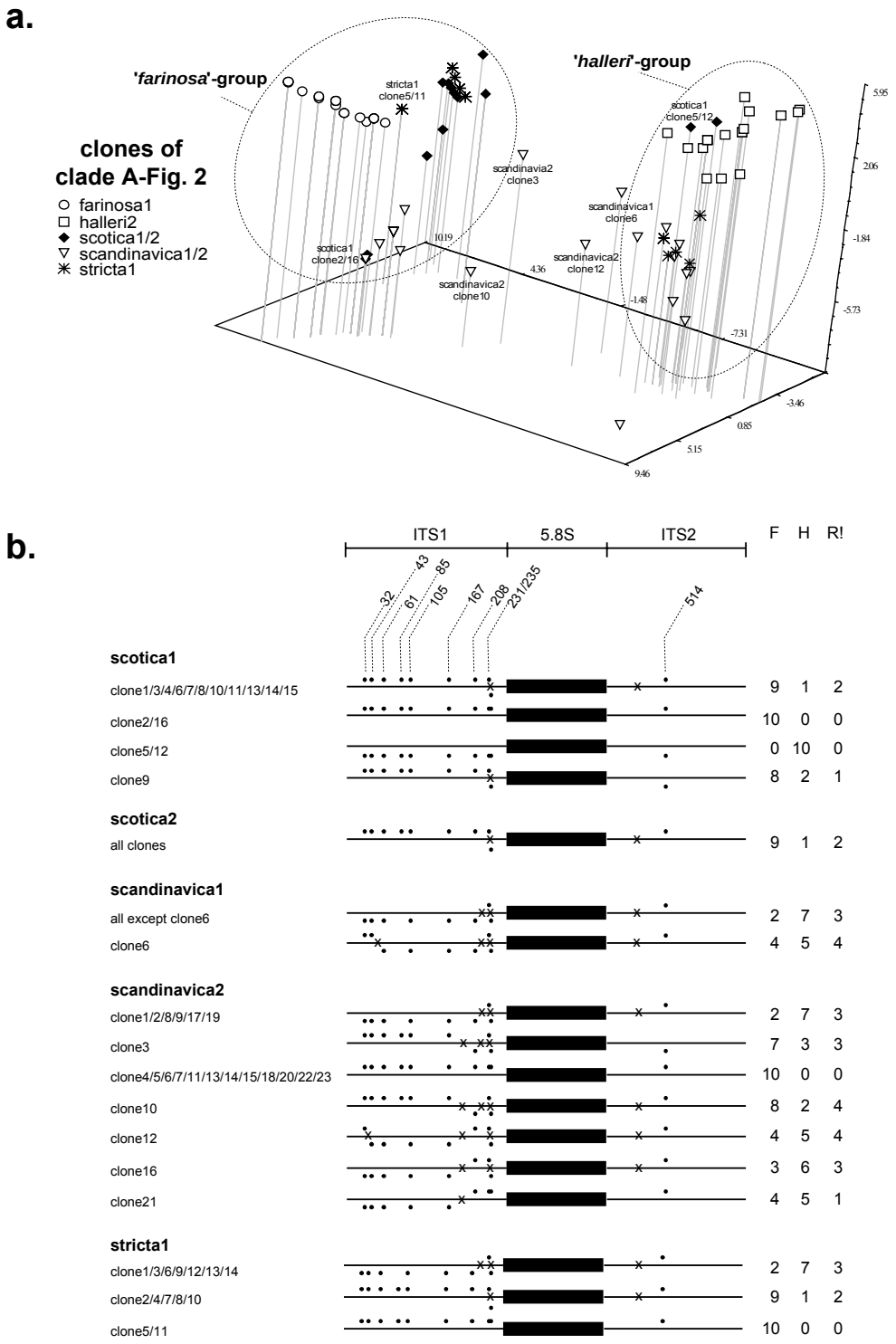


Figure 5. Analyses of variation among ITS clones. (a) Three-dimensional scatterplot for the first three axes obtained by PCoA of 91 cloned ITS sequences from accessions of clade A (cf. Fig. 2). Clones discussed in the text are labelled accordingly. (b) Recombination in cloned ITS sequences of *Primula scotica* (6x), *P. scandinavica* (8x) and accessions stricta1 of *P. stricta* (14x). Lines represent ITS1 and ITS2, filled boxes the 5.8S ribosomal gene. Site numbering follows the ITS alignment. The ten sites that differ between *P. farinosa* and *P. halleri* were mapped on diagrams of the cloned sequences (cf. Table 5). Filled circles placed above and below the lines denote synapomorphies with *P. farinosa* and *P. halleri*, respectively. A recombination event was inferred each time there was a switch (identified by a cross) from a nucleotide specific to *P. farinosa* to another specific to *P. halleri*, or vice-versa. F, number of synapomorphies shared with *P. farinosa*; H, number of synapomorphies shared with *P. halleri*; R!, minimum number of recombinations necessary to explain the observed nucleotides. Abbreviations of samples as in online Appendices 1 and 3.

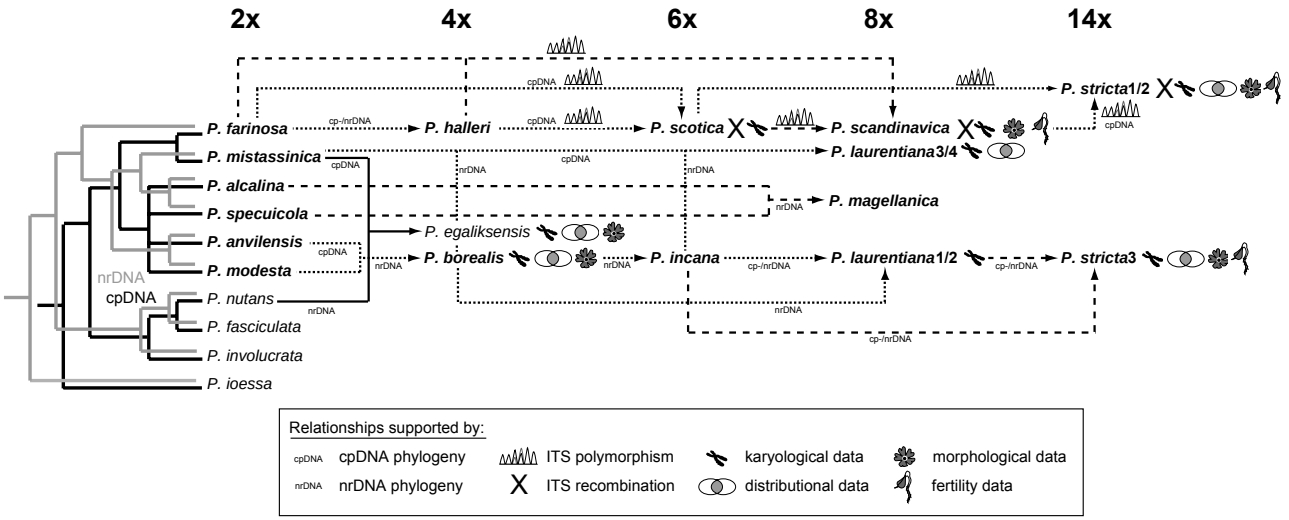


Figure 6. Patterns of reticulation inferred for *Primula* sect. *Aleuritia*. For diploid lineages (2x), phylogenetic relationships derived from cpDNA (black) and nrDNA (grey) data are represented by superimposed trees. For polyploid lineages (4x→14x), ancestor-descendant affinities are indicated by arrows and appositional symbols specify the type of data supporting the respective relationships. Species ascribed to *Primula* sect. *Aleuritia* are in bold. Polyphyletic species are abbreviated according Appendix 1.

(Fig. 5a). Different lines of evidence thus suggest an allopolyploid origin of the European *P. stricta* from hybridisation between a *P. scotica*- and a *P. scandinavica*-like ancestor and support the conclusion that *P. stricta* originated independently from different progenitor pairs on the two sides of the Atlantic (Fig. 6).

The mixed cloning experiment recovered 36% (33 out of 92) PCR-mediated recombinant clones. The fact 88% (30 out of 34), 71% (29 out of 41) and 86% (12 out of 14) recombinants were found in *P. scotica*, *P. scandinavica*, and sample stricta1 of *P. stricta*, respectively (Fig. 5b), suggests that a non-negligible proportion of recombinant clones in these polyploids do not result from PCR artefacts. This conclusion is corroborated by the finding that for a majority (79%) of the recombinant clones from the mixed cloning experiment only one recombination was necessary to explain the observed patterns of nucleotide composition (data not shown), while a minority (3%) of all the recombinant clones obtained from *P. scotica*, *P. scandinavica* and *P. stricta* were explained by a single recombination event (see scotica1 clone9 and scandinavica2 clone21; Fig. 5b).

Discussion

The use of multiple approaches, including classic phylogenetic and more recently developed consensus network methods, the identification of APS, the multivariate analysis of cloned ITS sequences, and the detection of recombination in ITS sequences, allowed us to gain new insights into the complex reticulate history of *Aleuritia*. Specifically, we were able to identify the putative parental lineages of the polyploids and compare our hypotheses with those previously proposed on the basis of caryological, morphological, and distributional evidence. It was also possible to

establish that most polyploids derived from hybridisation events and that some of them likely formed recurrently, conclusions that appear to be compatible with the secondary contact model (Stebbins, 1984, 1985) earlier invoked to explain speciation in *Aleuritia* (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). Below, we elucidate the origin of an intersectional hybrid involving a species of *Aleuritia* and summarise the evidence supporting the secondary contact model. A general theme emerging from our multi-faceted study is that each analytical method provided complementary insights that allowed us to disentangle different aspects of phylogenetic incongruence and, ultimately, reticulation in the complex.

Intersectional hybridisation

The North-American tetraploid *P. egaliksensis* (2n=40) was taxonomically ascribed to *Armerina* owing to its entire, petiolate leaves, narrow, elongate capsules, and the absence of farina (Kelso, 1991, 1992). The proposed hybrid origin of *P. egaliksensis* from the diploid, North- American *P. mistassinica* (2n=18; Fig. 3b) of *Aleuritia* and the diploid, circumboreal *P. nutans* (2n=22) of *Armerina* relied primarily on the additivity of chromosome numbers, the intermediacy of gland types, pollen sizes, colpi numbers and exine reticulation patterns, and the overlapping distribution of the mentioned diploids in north-western Canada, along the border with Alaska (Kelso, 1991, 1992). The present study corroborates the proposed origin of *P. egaliksensis*, while further allowing for the identification of the likely maternal (a *P. mistassinica*-like lineage) and paternal progenitors (a *P. nutans*-like lineage; Fig. 6), in agreement with recent cytogenetic investigations (Guggisberg *et al.*, 2008).

Support for the secondary contact model in Aleuritia

It has been repeatedly suggested that the *Aleuritia*-polyploid complex originated within the framework of the secondary contact model (Stebbins, 1984, 1985) through hybridisation events that involved differentiated populations of two widespread diploids, a *P. mistassinica*-like ancestor in America and a *P. farinosa*-like ancestor in Europe (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). The DNA-based data presented here are congruent with the hypothesis that cycles of habitat fragmentation and reconnection driven by glacial advancement and retreat during the Pleistocene promoted active speciation within the *Aleuritia* polyploid complex.

Our molecular data are compatible with previously proposed hybrid origins and putative parental lineages for the polyploids (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002), except for possible speciation via autopolyploidisation in the tetraploids *P. borealis* and *P. halleri* and in the octoploid *P. magellanica* (Fig. 6). For example,

they suggest an allopolyploid origin of *P. scotica* involving a *P. farinosa*- and a *P. halleri*-like ancestor (Fig. 6), adding further details to the general hypothesis of a hybrid origin for *P. scotica* inferred from the high levels of intra-population allelic variation revealed by allozyme markers (Glover and Abbott, 1995). Interestingly, *P. farinosa* currently occurs in northern England (Fig. 3b), and seeds probably belonging to *P. halleri* were discovered in Quaternary deposits of Cambridegshire, along with seeds attributed to *P. scotica* (Dovaston, 1956). Likewise, our data support the assumption that the octoploid *P. scandinavica* is presumably related with the hexaploid *P. scotica* and the diploid *P. farinosa*, for it shares the presence of dense farina on vegetative parts and the perforate pollen ornamentations with the former (Hultgård, 1990, 1993), while showing high levels of cross-fertility with the latter (Arnold and Richards, 1998). In contrast, phylogenetic evidence imply that the South-American octoploid *P. magellanica* might derive from diploid *P. alcalina*- and/or *P. specuicola*- like ancestors (Fig. 6), even though it morphologically resembles the two heavily farinose, North-American polyploids *P. incana* (6x) and *P. laurentiana* (8x; Richards, 2002).

The present study further corroborates the hypothesis that diploid *P. mistassinica*-like and *P. farinosa*-like ancestors played a role in the evolution of polyploid series on the American and European continents, respectively (Fig. 6). However, it also shows that additional diploid lineages, likely related to the extant *P. alcalina*, *P. anvilensis*, *P. modesta*, and *P. specuicola*, were involved in the origin of the American polyploids (Fig. 6). For example, it confirms earlier taxonomic investigations suggesting that the amphi-Beringian tetraploid *P. borealis* might be related to the diploid endemics *P. anvilensis* and *P. modesta* (Fig. 6), because of their overlapping distribution in the Alaskan Seward peninsula and Japan (Fig. 3b), respectively, and the similarity of the large conspicuous flowers and symmetrical umbel inflorescences of the tetraploid with the latter diploid (Kelso, 1991, 1992).

Finally, the likely independent origins of geographically separated accessions ascribed to *P. laurentiana* (8x) and *P. stricta* (14x) from different progenitors (Fig. 6) are congruent with the expectations of repeated speciation via secondary contacts. For instance, based on additivity of chromosome numbers, cross-fertility experiments and distributional patterns, two progenitor pairs were proposed for *P. stricta* (14x): the North-American *P. incana* (6x) and *P. laurentiana* (8x) and the European *P. scotica* (6x) and *P. scandinavica* (8x; Kelso, 1991, 1992; Arnold and Richards, 1998). Combined evidence from different analyses support polyphyletic origins for the amphi-Atlantic *P. stricta* and suggest that the North-American and European accessions of *P. stricta* originated independently from the two proposed progenitor pairs on the two sides of the Atlantic (Fig. 6).

Identifying causes of phylogenetic incongruence

The comparison between cpDNA and nrDNA phylogenies of *Aleuritia* (Figs. 1-2) revealed several cases of topological inconsistencies and the analysis of consensus networks computed from the simplified 70% BS consensus trees (Fig. 3a) suggested that the polyploids are mostly responsible for the observed incongruence (Fig. 3e). Considering the previously published hypotheses of hybrid origins for the *Aleuritia* polyploids, developed on the basis of caryological, morphological, and distributional evidence (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002), it seems reasonable to point to hybridisation as a cause of phylogenetic incongruence in *Aleuritia*. However, the evolutionary causes of phylogenetic incongruence may also include paralogy, lineage sorting, concerted evolution and recombination (Doyle, 1992; Maddison, 1997; Doyle and Davis, 1998; Wendel and Doyle, 1998; Álvarez and Wendel, 2003; Bailey *et al.*, 2003; Linder and Rieseberg, 2004). Below we illustrate selected examples of how different analytical methods induced us to favour reticulation as the major source of incongruence in the *Aleuritia* complex.

Phylogenetic incongruence can occur when gene duplication events precede cladogenetic divergence. In this case, discordance between cpDNA and nrDNA trees may be caused by the use of paralogous, rather than orthologous gene sequences for phylogenetic inference (Doyle, 1992; Maddison, 1997; Doyle and Davis, 1998; Álvarez and Wendel, 2003; Bailey *et al.*, 2003; Linder and Rieseberg, 2004). Barring extinction of gene copies, it is ideally necessary to sample all repeats to reduce the probability of erroneous phylogenetic comparisons. While our sampling of infra-specific ITS variation may not be exhaustive, the analysis of 179 ITS clones (excluding duplicates) allowed the additional detection of 23 additive nucleotides that had not been identified by direct sequencing of PCR products and that provided crucial details into the evolutionary history of *Aleuritia* (see Results). For example, ITS clones revealed 16 additional polymorphisms that link *P. halleri* (4x) with *P. scotica* (6x), solidifying the suggestion that the tetraploid contributed to the origin of the hexaploid (Fig. 6).

Both lineage sorting, the stochastic sorting of alleles following divergence from a polymorphic ancestor, and reticulation, the merging of differentiated genomes in a single species, may cause alleles from one species to be more closely related to alleles from a different species than to alleles from the same species (Doyle, 1992; Maddison, 1997; Wendel and Doyle, 1998), potentially causing phylogenetic incongruence. In the nrDNA and cpDNA consensus trees of *Aleuritia* (Figs. 1-2), all sequences from multiple accessions of the same species were monophyletic or members of polytomies, with the exception of sequences from the octoploid *P. laurentiana* and the 14-ploid *P. stricta*, which were polyphyletic in at least one tree and topologically incongruent between the two

trees (Fig. 3a). These results could be interpreted as evidence for either lineage sorting or reticulation in these two species. While our data are not fully conclusive on the causes of phylogenetic incongruence in the case of *P. laurentiana*, detailed analyses of ITS sequences proved more informative for *P. stricta*. More specifically, the observation that cloned ITS sequences of the European accession stricta1 cluster with either of the previously proposed parental lineages (i.e. *P. scotica* and *P. scandinavica*; Fig. 5a; Kelso, 1991, 1992; Arnold and Richards, 1998), reflecting recombination and different degrees of homogenisation towards the parental repeats (Fig. 5b), seems to favour the reticulation hypothesis for the observed incongruence between the cpDNA and nrDNA consensus trees (Fig. 3a).

Phylogenetic incongruence also arises when processes of genetic homogenisation in a multi-gene family of the nuclear genome cause the conversion of all copies in a given species to a single repeat type that differs in its relationships from those of the cpDNA topology (Doyle and Davis, 1998; Wendel and Doyle, 1998; Álvarez and Wendel, 2003). For example, in the tetraploid *P. egaliksensis*, all ITS sequences obtained directly from PCR products of eight individuals form a single clade with ITS sequences of the diploid *P. nutans* (Fig. 2), a topology that differs from the placement of *P. egaliksensis* accessions with the diploid *P. mistassinica* in the cpDNA tree (clade V, Fig. 1). Multivariate analyses of cloned PCR products showed that all ITS sequences from *P. egaliksensis* clustered with *P. nutans*, confirming that they had homogenised towards a single repeat-type (data not shown).

Finally, recombination between different repeat types, by producing sequences that possess a mixture of parental characteristics, often leads to conflicting hypotheses of relationships, resulting in a lack of phylogenetic resolution (Doyle and Davis, 1998; Wendel and Doyle, 1998; Álvarez and Wendel, 2003; Linder and Rieseberg, 2004). In the cpDNA and nrDNA consensus trees of *Aleuritia*, the relationships of the polyploid *P. scotica* are slightly inconsistent, for direct ITS sequences of this species form a poorly resolved clade with *P. stricta* (i.e. sample stricta2) and *P. farinosa* (clade A, Fig 2), while cpDNA sequences form a polytomy with *P. farinosa* and *P. halleri* (clade IV, Fig. 1). The analysis of cloned ITS sequences from *P. scotica* revealed that two clones each were completely homogenised towards the *P. farinosa*- (scotica1 clone2/16) and *P. halleri*-repeat type (scotica1 clone5/12), respectively, while 22 clones (excluding duplicates) showed evidence of recombination between sites differentiating *P. halleri* and *P. farinosa*, with partial homogenisation towards the *P. farinosa* type (Fig. 5b), thus suggesting that phylogenetic inconsistencies can be explained by recombination between parental ITS repeat types. The high rate of recombination detected in *P. scotica*, *P. scandinavica* and accession stricta1 of *P. stricta* may be the hallmark of ongoing processes of concerted evolution that ultimately lead to the homogenisation

of divergent ITS repeats in allopolyploids, for concerted evolution acts via intra- and interlocus recombination (Wendel *et al.*, 1995; Kovarík *et al.*, 2004; Kovarík *et al.*, 2005).

To summarise, phylogenetic incongruence may be suggestive of hybrid origins, but it is insufficient to eliminate other possible causes of discrepancies between phylogenies derived from differentially inherited genomes. Therefore, the use of different types of data (i.e. direct DNA sequences from maternally and biparentally inherited genomes and cloned sequences from a nuclear marker) and multiple analytical approaches (i.e. generation of cpDNA and nrDNA trees and their comparison via consensus networks, identification of APS, multivariate analysis of cloned ITS sequences, and analysis of recombination in ITS) is fundamental to achieve a comprehensive understanding of evolutionary relationships in polyploid complexes. In the study presented here, each line of evidence contributed to clarifying different details of the evolutionary history of *Aleuritia*. In combination with pre-existing knowledge from morphology, caryology and distribution, our multi-faceted approach allowed us to propose an integrative scenario of speciation via reticulation congruent with the secondary contact model in a polyploid, arctic-alpine group of plants. The next step will consist in investigating more nuclear regions to further distinguish between the different processes causing phylogenetic incongruence in our dataset and verify the conclusions presented in this paper.

Acknowledgments

For help with the collection of tissue samples we thank Torbjørn Alm, Alan Batten, Bruce Bennett, Sean Blaney, Christian Brochmann, Luc Brouillet, Edward Buyarski, Matthew Carlson, Mary Beth Cook, Reidar Elven, Pamela Eveleigh, Tupuna Kovanen, Wendy Maovlic, Michael Oldham, Carolyn Parker, Carl Roland, Peter Schönswetter, Heidi Solstad, Andreas Tribsch; all the colleagues listed in Appendix 1; and the following herbaria: ACAD, ALA, ALAJ, ALTA, AMNH, CAFB, CAN, CDFN, DAO, GB, HAM, ICEL, LKHD, MT, NFLD, NHIC, NSPM, O, QFA, S, SASK, SCFQ, SLU, SMI, TRH, TROM, TRTE, S, UBC, UME, UNB, UQTR, UPS, UWPG, WAT, WIN, WLK. We would also like to acknowledge John Richards and Sylvia Kelso for general guidance on *Primula*, and Jack Sullivan, Roberta Mason-Gamer and two anonymous reviewers for helpful comments on earlier versions of the manuscript. The Georges und Antoine Claraz-Schenkung and the Swiss Academy of Sciences funded AG's field trips. This study was initially facilitated by grant no. 3100-061674.00/1 from the Swiss National Science Foundation and funds from the Office for Equal Opportunity of the University of Zurich to EC.

References

Álvarez, I., and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29:417-434.

Arnold, E. S., and A. J. Richards. 1998. On the occurrence of unilateral incompatibility in *Primula* section *Aleuritia* Duby and the origin of *Primula scotica* Hook. Botanical Journal of the Linnean Society 128:359-368.

Arnold, M. L. 1997. Natural Hybridization and Evolution. Oxford University Press, Oxford, UK.

Arnold, M. L. 2006. Evolution Through Genetic Exchange. Oxford University Press, New York, NY, USA.

Bailey, C. D., T. G. Carr, S. A. Harris, and C. E. Hughes. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Molecular Phylogenetics and Evolution 29:435-455.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell, and M. J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82:247-277.

Baum, D. A., R. L. Small, and J. F. Wendel. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Systematic Biology 47:181-207.

Bretagnolle, F., F. Felber, F. G. Calame, and P. Küpfer. 1998. La polyploïdie chez les plantes. Botanica Helvetica 108:5-37.

Brochmann, C., Q.-Y. Xiang, S. J. Brunsfeld, D. E. Soltis, and P. S. Soltis. 1998. Molecular evidence for polyploid origins in S*axifraga* (Saxifragaceae): the narrow arctic endemic *S. svalbardensis* and its widespread allies. American Journal of Botany 85:135-143.

Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. Biological Journal of the Linnean Society 82:521-536.

Bruun, H. G. 1932. Cytological studies in *Primula*, with special reference to the relation between the karyology and taxonomy of the genus. Symbolae Botanicae Upsalienses 1:1-239.

Brysting, A. K., B. Oxelman, K. T. Huber, V. Moulton, and C. Brochmann. 2007. Untangling complex histories of genome merging in high polyploids. Systematic Biology 53:467-476.

Buckler, E. S., IV, A. Ippolito, and T. P. Holtsford. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821-832.

Campbell, C. S., M. F. Wojciechowski, B. G. Baldwin, L. A. Alice, and M. J. Donoghue. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Molecular Biology and Evolution 14:81-90.

Carlson, M. L., S. D. Gisler, and S. Kelso. 2008. The role of reproductive assurance in the Arctic: a comparative study of a homostylous and distylous species pair. Arctic, Antarctic, and Alpine Research 40:39-47.

Conti, E., E. Suring, D. Boyd, J. Jorgensen, J. Grant, and S. Kelso. 2000. Phylogenetic relationships and character evolution in *Primula* L.: the usefulness of ITS sequence data. Plant Biosystems 134:385-392.

Corriveau, J. L., and A. W. Coleman. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperms species. American Journal of Botany 75:1443-1458.

Demesure, B., N. Sodzi, and R. J. Petit. 1995. A set of universal primers for amplificaton of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology 4:129-131.

Devos, N., S.-H. Oh, O. Raspé, A.-L. Jacquemart, and P. S. Manos. 2005. Nuclear ribosomal DNA sequence variation and evolution of spotted marsh-orchids (*Dactylorhiza maculata* group). Molecular Phylogenetics and Evolution 36:568-580.

Devos, N., O. Raspé, S.-H. Oh, D. Tyteca, and A.-L. Jacquemart. 2006. The evolution of *Dactylorhiza* (Orchidaceae) allotetraploid complex: insights from nrDNA sequences and cpDNA PCR-RFLP data. Molecular Phylogenetics and Evolution 38:767-778.

Dovaston, H. F. 1956. *Primula scotica* Hook., a relict species in Scotland. Notes from the Royal Botanical Garden of Edinburgh 221:289-291.

Doyle, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17:144-163.

Doyle, J. J., and J. I. Davis. 1998. Homology in molecular phylogenetics: a parsimony perspective. Pages 101-131 *in* Molecular Systematics of Plants II: DNA Sequencing (D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds.). Kluwer Academic Publishers, Norwell, MA, USA.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the boostrap. Evolution 17:223-263.

Fuertes Aguilar, J., and G. Nieto Feliner. 2003. Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). Molecular Phylogenetics and Evolution 28:430-447.

Glover, B. J., and R. J. Abbott. 1995. Low genetic diversity in the Scottish endemic *Primula scotica* Hook. New Phytologist 129:147-153.

Gower, J. C. 1966. Some distance properties of latent root and vector used in multivariate analysis. Biometrika 53:325-338.

Grant, V. 1981. Plant Speciation. Columbia University Press, New York, NY, USA.

Guggisberg, A., F. Bretagnolle, and G. Mansion. 2006a. Allopolyploid origin of the Mediterranean endemic, *Centaurium bianoris* (Gentianaceae), inferred by molecular markers. Systematic Botany 31:368-379.

Guggisberg, A., G. Mansion, S. Kelso, and E. Conti. 2006b. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*. New Phytologist 171:617-632.

Guggisberg, A., C. Baroux, U. Grossniklaus, and E. Conti. 2008. Genomic origin and organisation of the allopolyploid *Primula egaliksensis* investigated by *in situ* hybridisation. Annals of Botany doi:10.1093/aob/mcn026.

Henry, Y., M. Bedhomme, and G. Blanc. 2006. History, protohistory and prehistory of the *Arabidopsis thaliana* chromosome complement. Trends in Plant Science 11:267-273.

Holland, B., and V. Moulton. 2003. Consensus networks: a method for visualising incompatibilities in collections of trees. Pages 165-176 *in* Algorithms in Bioinformatics: Proceedings of the Third International Workshop, WABI 2003 (G. Benson, and R. Page, eds.). Springer Verlag, Berlin, Germany.

Holland, B. R., F. Delsuc, and V. Moulton. 2005. Visualizing conflicting evolutionary hypothesis in large collections of trees: using consensus networks to study the origins of placentals and hexapods. Systematic Biology 54:66-76.

Hoot, S. B., N. S. Napier, and W. C. Taylor. 2004. Revealing unknown or extinct lineages within *Isoëtes* (Isoëtaceae) using DNA sequences from hybrids. American Journal of Botany 91:899-904.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.

Huelsenbeck, J. P., B. Larget, R. E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Systematic Biology 51:673-688.

Hughes, C. E., C. D. Bailey, and S. A. Harris. 2002. Divergent and reticulate species relationships in *Leucaena* (Fabaceae) inferred from multiple data sources: insights into polyploid origins and nrDNA polymorphism. American Journal of Botany 89:1057-1073.

Hultgård, U.-M. 1990. Polyploidy and differentiation in N European populations of *Primula* subgenus *Aleuritia*. Sommerfeltia 11:117-135.

Hultgård, U.-M. 1993. *Primula scandinavica* and *P. stricta* - patterns of distribution, variation, reproductive strategies and migrations. Opera Botanica 121:35-43.

Huson, D. 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14:68-73.

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Systematic Biology 23:254-267.

Kelso, S. 1991. Taxonomy of *Primula* sects. *Aleuritia* and *Armerina* in North America. Rhodora 93:67-99.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. Arctic and Alpine Research 24:82-87.

Kovarík, A., R. Matyásek, K. Y. Lim, K. Skalická, B. Koukalová, S. Knapp, M. Chase, and A. R. Leitch. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. Biological Journal of the Linnean Society 82:615-625.

Kovarík, A., J. C. Pires, A. R. Leitch, K. Y. Lim, A. M. Sherwood, R. Matyásek, J. Rocca, D. E. Soltis, and P. S. Soltis. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. Genetics 169:931-944.

Levin, D. A. 2002. The Role of Chromosomal Change in Plant Evolution. Oxford University Press, New York, NY, USA.

Linder, C. R., and L. H. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. American Journal of Botany 91:1700-1708.

Long, E. O., and I. B. Dawid. 1980. Repeated genes in eukaryotes. Annual Review of Biochemistry 49:727-764.

Maddison, W. P. 1997. Gene trees in species trees. Systematic Biology 46:523-536.

Mansion, G., L. Zeltner, and F. Bretagnolle. 2005. Phylogenetic patterns and polyploid evolution within the Mediterranean genus *Centaurium* (Gentianaceae - Chironieae). Taxon 54:931-950.

Mason-Gamer, R. 2004. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. Systematic Biology 53:25-37.

Mast, A. R., S. Kelso, and E. Conti. 2006. Are any primroses (*Primula*) primitively monomorphic? New Phytologist 171:605-616.

Mast, A. R., S. Kelso, A. J. Richards, D. J. Lang, D. M. S. Feller, and E. Conti. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. International Journal of Plant Sciences 162:1381-1400.

Masterson, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264:421-424.

McBreen, K., and P. J. Lockhart. 2006. Reconstructing reticulate evolutionary histories of plants. Trends in Plant Science 11:398-404.

McDade, L. A. 1995. Hybridization and phylogenetics. Pages 305-331 *in* Experimental and Molecular Approaches to Plant Biosystematics (P. C. Hoch, and A. G. Stephenson, eds.). Missouri Botanical Garden, St.-Louis, MI, USA.

Morrison, D. A. 2005. Networks in phylogenetic analysis: new tools for population biology. International Journal of Parasitology 35:567-582.

Nieto Feliner, G., and J. A. Rosselló. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution 44:911-919.

Noyes, R. D. 2006. Intraspecific nuclear ribosomal DNA divergence and reticulation in sexual diploid *Erigeron strigosus* (Asteraceae). American Journal of Botany 93:470-479.

Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53:47-67.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Annual Review of Ecology and Systematics 34:401-437.

Oxelman, B., M. Liden, and D. Berglund. 1997. Chloroplast *rps*16 intron phylogeny of the tribe Sileneae (Caryophyllaceae). Plant Systematics and Evolution 206:393-410.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Posada, D., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. Trends in Ecology and Evolution 16:37-45.

Rambaud, A., and A. J. Drummond. 2007. Tracer, version 1.4. Available from http://beast.bio.ed.ac.uk/Tracer.

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29:467-501.

Richards, J. 2002. *Primula*. B. T. Batsford, London, UK.

Rieseberg, L. H. 1997. Hybrid origins of plant species. Annual Review of Ecology and Systematics 28:359-89.

Rieseberg, L. H., and N. C. Ellstrand. 1993. What can molecular and morphological markers tell us about plant hybridization? Critical Review in Plant Sciences 12:213-241.

Rieseberg, L. H., and S. E. Carney. 1998. Plant hybridization. New Phytologist 140:599-624.

Rohlf, F. J. 2000. NTSYSpc Numerical Taxonomy and Multivariate Analysis System, version 2.1. Applied Biostatistics Inc.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Sang, T., D. J. Crawford, and T. F. Stuessy. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Science, USA 92:6813-6817.

Sang, T., D. J. Crawford, and T. F. Stuessy. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). American Journal of Botany 84:1120-1136.

Simmons, M. P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology 49:369-381.

Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany 17:145-170.

Small, R. L., J. A. Ryburn, R. C. Cronn, T. Seelanan, and J. F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. American Journal of Botany 85:1301-1315.

Smedmark, J. E. E., T. Eriksson, R. C. Evans, and C. S. Campbell. 2003. Ancient allopolyploid speciation in *Geinae* (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. Systematic Biology 52:374-385.

Soltis, D. E., and P. S. Soltis. 1993. Molecular data and dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243-273.

Soltis, D. E., and R. K. Kuzoff. 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). Evolution 49:727-742.

Soltis, D. E., and P. S. Soltis. 1999. Polyploidy: recurrent formation and genome evolution. Trends in Ecology and Evolution 14:348-352.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2003. Advances in the study of polyploidy since *Plant Speciation*. New Phytologist 161:173-191.

Stebbins, G. L. 1950. Variation and Evolution in Plants. Columbia University Press, New York, NY, USA.

Stebbins, G. L. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. Botanica Helvetica 94:1-13.

Stebbins, G. L. 1985. Polyploidy, hybridization, and the invasion of new habitats. Annals of the Missouri Botanical Garden 72:824-832.

Steen, S. W., L. Gielly, P. Taberlet, and C. Brochmann. 2000. Same parental species, but different taxa: molecular evidence for hybrid origins of the rare endemics *Saxifraga opdalensis* and *S. svalbardensis* (Saxifragaceae). Botanical Journal of the Linnean Society 132:153-164.

Swofford, D. L. 1999. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), version 4. Sinauer Associates.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17:1105-1109.

Trift, I., M. Källersjö, and A. A. Anderberg. 2002. The monophyly of *Primula* (Primulaceae) evaluated by analysis of sequences from the chloroplast gene *rbc*L. Systematic Botany 27:396-407.

Vision, T., D. G. Brown, and S. D. Tanksley. 2000. The origins of genomic duplications in *Arabidopsis*. Science 290:2114-2117.

Vogelmann, H. W. 1956. A Biosystematic Study of *Primula mistassinica* Michx. University of Michigan, Ann Arbor, MI, USA.

Vriesendorp, B., and F. T. Bakker. 2005. Reconstructing patterns of reticulate evolution in angiosperms: what can we do? Taxon 54:593-604.

Wagner, A., N. Blackstone, P. Artwright, M. Dick, B. Misof, P. Snow, G. P. Wagner, J. Bartels, M. Murtha, and J. Pendleton. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Systematic Biology 43:250-261.

Wendel, J. F., and J. J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. Pages 265-296 *in* Molecular Systematics of Plants II: DNA Sequencing (D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds.). Kluwer Academic Publishers, Norwell, MA, USA.

Wendel, J. F., A. Schnabel, and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proceedings of the National Academy of Science, USA 92:280-284.

Young, N. D., and J. Healy. 2003. GapCoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4.

Zhang, L.-B., and J. W. Kadereit. 2004. Classification of *Primula* sect. *Auricula* (Primulaceae) based on two molecular data sets (ITS, AFLP), morphology and geographical distribution. Botanical Journal of the Linnean Society 146:1-26.

Zhang, L.-B., H. P. Comes, and J. W. Kadereit. 2004. The temporal course of Quaternary diversification in the European high mountain endemic *Primula* sect. *Auricula* (Primulaceae). International Journal of Plant Sciences 165:191-207.

Appendix 1. Species of *Primula* sampled for the present study. Data are presented in the following sequence: Taxon, Section, OTU abbreviation (in bold), Population location, Voucher information (italicised), GenBank accession numbers for *rpl*16 intron, *rps*16 intron, *trn*L intron, *trn*L-*trn*F spacer, *trn*T-*trn*L spacer, *trn*D-*trn*T region, and ITS, respectively (references for clones are indicated in parentheses). GenBank accession numbers with literature references are for sequences obtained by other authors, whereas those that are underlined are new. na, not available.

Primula alcalina Cholewa & Henderson, sect. *Aleuritia* subsect. *Aleuritia*, **alcalina**, USA, Idaho, Targhee National Forest, *Lehman s.n.* (Z), AF402489 (Mast *et al.*, 2001), DQ379908, AF402370 (Mast *et al.*, 2001), DQ379731, DQ379792, DQ994079, DQ993750 (clones 1-22: DQ993803-DQ993824)

Primula anvilensis Kelso, sect. *Aleuritia* subsect. *Aleuritia*, **anvilensis**, USA, Alaska, Nome, *Guggisberg & Mansion 200703-1* (Z), DQ379852, DQ379911, DQ379733, DQ379733, DQ379794, DQ994080, DQ993731

Primula borealis Duby, sect. *Aleuritia* subsect. *Aleuritia*, **borealis1**, USA, Alaska, Unalakleet, *Parker 7537* (ALA), AF402488 (Mast *et al.*, 2001), DQ379913, AF402369 (Mast *et al.*, 2001), DQ379735, DQ379796, DQ994081, DQ993732; **borealis2**, USA, Alaska, Kotzebue, *Elven & Solstad S03-11* (O), DQ379853, DQ379912, DQ379736, DQ379736, DQ379797, DQ994082, DQ993733 (clones 1-16: DQ993843-DQ993858)

Primula egaliksensis Wormsk. ex Hornem., sect. *Armerina*, **egaliksensis1**, Canada, Yukon Territory, Abraham Mt., *Elven & Solstad SUP03-352* (O), DQ379875, DQ379936, DQ379759, DQ379759, DQ379817, DQ994120, DQ993755; **egaliksensis2**, Canada, Newfoundland, St.-Barbe, *Guggisberg & Mansion 250604-8* (Z), DQ379876, DQ379937, DQ379763, DQ379763, DQ379823, DQ994121, DQ993756; **egaliksensis3**, USA, Alaska, Goodnews Bay, *Parker 160604-2* (Z), DQ379877, DQ379940, DQ379762, DQ379762, DQ379826, DQ994122, DQ993757; **egaliksensis4**, USA, Colorado, South Park, *Kelso 060703-2* (Z), DQ379878, DQ379938, DQ379764, DQ379764, DQ379824, DQ994123, DQ993758; **egaliksensis5**, USA, Alaska, Glacier Bay NP, Gustavus, *Bosworth 160603-1* (Z), DQ379879, DQ379939, DQ379765, DQ379765, DQ379825, DQ994124, DQ993759; **egaliksensis6**, USA, Alaska, Wrangell-St.Elias NP, Jake Lake, *Guggisberg & Mansion 230603-1* (Z), DQ379880, DQ379941, DQ379761, DQ379761, DQ379827, DQ994125, DQ993760; **egaliksensis7**, Canada, Newfoundland, Port au Choix Peninsula, *Guggisberg & Mansion 240604-6b* (Z), DQ379881, DQ379942, DQ379760, DQ379760, DQ379828, DQ994126, DQ993761 (clones 1-20: DQ993783-DQ993802); **egaliksensis8**, USA, Alaska, Cape Krusenstern, *Parker 10392* (ALA), DQ379882, DQ379943, DQ379766, DQ379766, DQ379829, DQ994127, DQ993762

Primula fasciculata Balf. f. & Ward, sect. *Armerina*, **fasciculata**, horticultural origin, *McBeath s.n.* (Z), DQ379891, DQ379953, DQ379776, DQ379776, DQ379839, DQ994137, DQ993773

Primula farinosa L., sect. *Aleuritia* subsect. *Aleuritia*, **farinosa1**, Italy, South Tirol, Schlüsseljochs, *Schönswetter & Tribsch 100703-3* (Z), DQ379845, DQ379897, DQ379720, DQ379720, DQ379781, DQ994098, DQ993712 (clones 1-24: DQ993939-DQ993962); **farinosa2**, Spain, Aragon, Central Pyrenees, Macizo La Maladeta, *Schneeweiss & Schönswetter 230703-5* (Z), DQ379846, DQ379898, DQ379722, DQ379722, DQ379784, DQ994099, DQ993713; **farinosa3**, Switzerland, Bern, Faulhornstock, *Wiedermann 230803-1* (Z), DQ379848, DQ379899, DQ379723, DQ379723, DQ379785, DQ994100, DQ993714; **farinosa4**, UK, Teesdale, Sand Syke, *Richards s.n.* (Z), DQ379849, DQ379900, DQ379716, DQ379716, DQ379787, DQ994101, DQ993715; **farinosa5**, UK , Cumbria, Gait Barrow, *Richards s.n.* (Z), AF402474 (Mast *et al.*, 2001), DQ379901, AF402356 (Mast *et al.*, 2001), DQ379725, DQ379786, DQ994102, DQ993716; **farinosa6**, France, Aquitaine, Atlantic Pyrenees, El Portalet, *Schneeweiss & Schönswetter 250703-1* (Z), DQ379842, DQ379893, DQ379724, DQ379724, DQ379782, DQ994103, DQ993717; **farinosa7**, Italy, Cuneo, Cottian Alps, Col de Sampeyre, *Wiedermann 160703-1* (Z), DQ379843, DQ379894, DQ379721, DQ379721, DQ379783, DQ994104, DQ993718; **farinosa8**, Slowakia, W Carpathians, Vel'ka Fatra Mts., *Ronikier & Ronikier 020503-1* (Z), DQ379844, DQ379895, DQ379718, DQ379718, DQ379779, DQ994105, DQ993719; **farinosa9**, Poland, W Carpathians, Beskid Sadecki massif, *Ronikier & Ronikier 120903-1* (na), DQ379847, DQ379896, DQ379719, DQ379719, DQ379780, DQ994106, DQ993720

Primula halleri Gmel., sect. *Aleuritia* subsect. *Aleuritia*, **halleri1**, Austria, Hohe Tauern, Ankogelgruppe, *Schönswetter & Tribsch 310503-1* (Z), DQ379850, DQ379903, DQ379726, DQ379726, DQ379789, DQ994108, DQ993721; **halleri2**, Italy, South Tirol, Schlüsseljochs, *Schönswetter & Tribsch 100703-3* (Z), DQ994140, DQ994141, DQ994143, DQ994143, DQ994142, DQ994107, DQ993722 (clones 1-21: DQ993963-DQ993983)

Primula incana M. E. Jones, sect. *Aleuritia* subsect. *Aleuritia*, **incana1**, Canada, Yukon Territory, Whitehorse, Takhini Salt Flats, *Guggisberg & Mansion 020703-1* (Z), DQ379855, DQ379915, DQ379742, DQ379742, DQ379799, DQ994092, DQ993734 (clones 1-18: DQ993859-DQ993876); **incana2**, USA, Alaska, Fairbanks, *Guggisberg & Mansion 080703-1-1* (Z), DQ379854, DQ379914, DQ379740, DQ379740, DQ379798, DQ994093, DQ993735; **incana3**, Canada, Yukon Territory, Meadow Lake, *Guggisberg & Mansion 040703-3* (Z), DQ379856, DQ379916, DQ379741, DQ379741, DQ379800, DQ994094, DQ993736

Primula involucrata Wallich, sect. *Armerina*, **involucrata**, horticultural origin, *Carrie s.n.* (Z), DQ379892, DQ379954, DQ379777, DQ379777, DQ379840, DQ994138, DQ993772

Primula ioessa Smith, sect. *Sikkimensis*, **ioessa**, horticultural origin, *Tromsoy Botanical Garden 92-1052* (na), AF402501 (Mast *et al.*, 2001), DQ379955, AF402381 (Mast *et al.*, 2001), DQ379778, DQ379841, DQ994139, DQ993774

Primula laurentiana Fernald, sect. *Aleuritia* subsect. *Aleuritia*, **laurentiana1**, Canada, Newfoundland, Green Point, *Guggisberg & Mansion 230604-3* (Z), DQ379857, DQ379917, DQ379744, DQ379744, DQ379803, DQ994095, DQ993737 (clones 1-22: DQ993877-DQ993898); **laurentiana2**, Canada, Newfoundland, Reefs Harbour, *Guggisberg & Mansion 260604-1* (Z), DQ379858, DQ379918, DQ379743, DQ379743, DQ379804, DQ994096, DQ993738; **laurentiana3**, Canada, Quebec, St.-Lawrence River, St.-Simon, *Guggisberg & Mansion 130604-2* (Z), DQ379859, DQ379920, DQ379747, DQ379747, DQ379802, DQ994088, DQ993739; **laurentiana4**, Canada, New Brunswick, Fundy Bay, *Guggisberg & Mansion 180604-1* (Z), DQ379860, DQ379919, DQ379746, DQ379746, DQ379801, DQ994089, DQ993740 (clones 1-19: DQ993899-DQ993917)

Primula magellanica Lehm., sect. *Aleuritia* subsect. *Aleuritia*, **magellanica1**, Argentina, Tierra del Fuego, *Richards s.n.* (na), DQ379865, DQ379922, DQ379739, DQ379739, DQ379812, DQ994083, DQ993752; **magellanica2**, E Falkland Is., Stanley aiport, *Thompson & Woods 251103-1* (na), DQ379864, DQ379921, DQ379737, DQ379737, DQ379810, DQ994084, DQ993753 (clones 1-9: DQ993825-DQ993833); **magellanica3**, horticultural origin, *Lever s.n.* (na), AF402482 (Mast *et al.*, 2001), DQ379923, AF402364 (Mast *et al.*, 2001), DQ379738, DQ379811, DQ994085, DQ993754

Primula misstassinica Michaux, sect. *Aleuritia* subsect. *Aleuritia*, **misstassinica1**, Canada, Quebec, St.-Lawrence River, Beaumont, *Guggisberg & Mansion 130604-1* (Z), DQ379869, DQ379931, DQ379751, DQ379751, DQ379818, DQ994111, DQ993741; **misstassinica2**, Canada, Newfoundland, Fischells Brook, *Guggisberg & Mansion 210604-1* (Z), DQ379870, DQ379927, DQ379753, DQ379753, DQ379813, DQ994112, DQ993742; **misstassinica3**, Canada, Ontario, Lake Huron, Manitoulin I., *Guggisberg & Mansion 100604-2* (Z), DQ379871, DQ379930, DQ379758, DQ379758, DQ379816, DQ994113, DQ993743; **misstassinica4**, Canada, Ontario, Lake Huron, Port Elgin, *Guggisberg & Mansion 090604-1* (Z), DQ379872, DQ379929, DQ379757, DQ379757, DQ379815, DQ994114, DQ993744; **misstassinica5**, Canada, Newfoundland, Reefs Harbour, *Guggisberg & Mansion 260604-1* (Z), DQ379873, DQ379934, DQ379754, DQ379754, DQ379821, DQ994115, DQ993745; **misstassinica6**, Canada, New Brunswick, Rothesay, *Guggisberg & Mansion 170604-2* (Z), DQ379874, DQ379932, DQ379752, DQ379752, DQ379819, DQ994116, DQ993746; **mistassinica7**, Canada, British Columbia, Boya Lake, *Guggisberg & Mansion 300603-1* (Z), DQ379866, DQ379928, DQ379755, DQ379755, DQ379814, DQ994117, DQ993747; **mistassinica8**, USA, Wisconsin, Newport State Park, *Anderson 210503-1a* (Z), DQ379867, DQ379933, DQ379750, DQ379750, DQ379820, DQ994118, DQ993748; **mistassinica9**, Canada, Alberta, Bow Valley; *Eveleigh 300603-3* (Z), DQ379868, DQ379935, DQ379756, DQ379756, DQ379822, DQ994119, DQ993749 (clones 1-9: DQ993834-DQ993842)

Primula modesta Bisset & Moore, sect. *Aleuritia* subsect. *Aleuritia*, **modesta**, horticultural origin, *Richards s.n.* (Z), AF402490 (Mast *et al.*, 2001), DQ379910, AF402371 (Mast *et al.*, 2001), DQ379734, DQ379795, DQ994077, DQ993730

Primula nutans Georgi, sect. *Armerina*, **nutans1**, Mongolia, Muren Hovsgol aimag, Delgermuren River valley, *Oyumaa 090703-1* (Z), DQ379883, DQ379944, DQ379767, DQ379767, DQ379830, DQ994128, DQ993763; **nutans2**, Russia, Altai Republic, valley of Tarkhata River, *Tribsch 300703-1* (Z), DQ379886, DQ379949, DQ379773, DQ379773, DQ379834, DQ994129, DQ993764; **nutans3**, Mongolia, Muren Hovsgol aimag, Hatgal Huzuuvch, *Oyumaa 180703-1* (Z), DQ379889, DQ379950, DQ379772, DQ379772, DQ379836, DQ994130, DQ993765; **nutans4**, Russia, Irkutsk region, Lake Baikal, *Kovtonyuk s.n.* (na), DQ379890, DQ379951, DQ379774, DQ379774, DQ379838, DQ994131, DQ993766; **nutans5**, USA, Alaska, Northway, *Guggisberg & Mansion 240603-2* (Z), DQ379884, DQ379946, DQ379769, DQ379769, DQ379831, DQ994132, DQ993767; **nutans6**, USA, Alaska, Nome, *Guggisberg & Mansion 220703-6* (Z), DQ379885, DQ379945, DQ379768, DQ379768,

DQ379832, DQ994133, DQ993768; **nutans7**, Canada, Yukon Territory, Little Braeburn Lake, *Guggisberg & Mansion 030703-2* (Z), DQ379887, DQ379947, DQ379770, DQ379770, DQ379833, DQ994134, DQ993769 (clones 1-7: DQ993775-DQ993782); **nutans8**, USA, Alaska, Goodnews Bay, *Parker 150604-5* (Z), DQ379888, DQ379948, DQ379771, DQ379771, DQ379835, DQ994135, DQ993770; **nutans9**, USA, Alaska, Unalakleet, *Parker 6966* (ALA), AF402494 (Mast *et al.*, 2001), DQ379952, AF 402374 (Mast *et al.*, 2001), DQ379775, DQ379837, DQ994136, DQ993771

Primula scandinavica (Bruun) Bruun, sect. *Aleuritia* subsect. *Aleuritia*, **scandinavica1**, Norway, Tromsø, *Alsos & Tveraabak 250803-1* (Z), DQ379861, DQ379906, DQ379729, DQ379729, DQ379805, DQ994090, DQ993725 (clones 1-18: DQ994022-DQ994039); **scandinavica2**, Norway, Bodø, *Richards s.n.* (Z), AF402483 (Mast *et al.*, 2001), DQ379907, AF402365 (Mast *et al.*, 2001), DQ379730, DQ379806, DQ994091, DQ993726 (clones 1-23: DQ994040-DQ994062)

Primula scotica Hooker, sect. *Aleuritia* subsect. *Aleuritia*, **scotica1**, horticultural origin, *Tromsoy Botanical Garden s.n.* (na), DQ379851, DQ379904, DQ379727, DQ379727, DQ379791, DQ994109, DQ993723 (clones 1-16: DQ993984-DQ993999); **scotica2**, horticultural origin, *Lever s.n.* (na), AF402476 (Mast *et al.*, 2001), DQ379905, AF402358 (Mast *et al.*, 2001), DQ379728, DQ379790, DQ994110, DQ993724 (clones 1-22: DQ994000-DQ994021)

Primula specuicola Rydb., sect. *Aleuritia* subsect. *Aleuritia*, **specuicola**, USA, Utah, Negro Bill Canyon, *Kelso 00-59* (COCO), AF402487 (Mast *et al.*, 2001), DQ379909, AF402368 (Mast *et al.*, 2001), DQ379732, DQ379793, DQ994078, DQ993751

Primula stricta Hornem., sect. *Aleuritia* subsect. *Aleuritia*, **stricta1**, horticultural origin, *Tromsoy Botanical Garden s.n.* (na), DQ379863, DQ379925, DQ379748, DQ379748, DQ379808, DQ994086, DQ993727 (clones 1-14: DQ994063-DQ994076); **stricta2**, horticultural origin, *Mattingley s.n.* (na), AF402484 (Mast *et al.*, 2001), DQ379926, AF402366 (Mast *et al.*, 2001), DQ379749, DQ379809, DQ994087, DQ993728; **stricta3**, Canada, Newfoundland, L'Anse-aux-Meadows, *Guggisberg & Mansion 250604-2* (Z), DQ379862, DQ379924, DQ379745, DQ379745, DQ379807, DQ994097, DQ993729 (clones 1-21: DQ993918-DQ993938)

Appendix 2. Eighty-eight variable sites within *Primula* sect. *Aleuritia* detected from direct sequences of ITS amplicons. Forty-three sites are polymorphic: the 94 polymorphisms are highlighted in grey (IUPAC ambiguity codes: M=A or C, R=A or G, W=A or T, S=C or G, Y=C or T, K=G or T). Site numbering follows the ITS alignment. Sites identical to the first sequence are represented by dots. Sites belonging to the 5.8S gene and separating ITS1 from ITS2 are italicised (i.e. sites 378, 391, and 413). Accessions followed by an asterisk were cloned (see Appendix 3). Accession abbreviations as in Appendix 1.

Accession	111	112	113	121	132	138	143	144	150	157	158	161	174	178	184	185	192	1102	1105	1106	1107	1108	1109	1116	1123	1125	1126	1127	1128	1129	1130
mistassinica1	T	G	A	T	C	C	A	C	T	T	G	T	G	G	C	G	C	G	C	A	C	G	C	C	G	G	C	G	T	G	T
mistassinica2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
mistassinica4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**Y**	.	.	.	.	.	.	.	.	.	.
mistassinica5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**Y**	.	.	.	.	.	.	.	.	.	.
mistassinica7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica9*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**Y**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
specuicola	.	A	.	.	.	.	.	.	.	G	.	C	A	.	.	.	T	.	.	.	A	.	.	.	.	.	T	.	.	.	.
alcalina*	.	A	.	.	.	.	.	T	C	G	.	C	A	.	.	.	T	.	.	.	.	.	.	.	T	.	T	.	.	.	.
magellanica1	.	A	G	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.
magellanica2*	.	A	G	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.
magellanica3	.	A	G	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.
stricta3*	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
modesta	.	A	.	.	.	.	.	.	.	G	.	C	.	C	.	.	T	.	.	G	.	.	.	A	C	.	.	-	-	.	.
anvilensis	C	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
borealis1	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
borealis2*	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
incana1*	A	A	.	.	.	.	.	.	.	G	**R**	C	.	.	.	.	T	.	.	G	.	.	**Y**	.	.	.	.	-	-	.	.
incana2	A	A	.	.	.	.	.	.	.	G	**R**	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
incana3	A	A	.	.	.	.	.	.	.	G	**R**	C	.	.	.	.	T	.	.	G	.	.	**Y**	.	.	.	.	-	-	.	.
laurentiana1*	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
laurentiana2	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
laurentiana3	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
laurentiana4*	A	A	.	.	.	.	.	.	.	G	.	C	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	-	-	.	.
farinosa1*	.	A	.	C	T	**Y**	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	**R**	**W**	.	.	.	.	-	-	A	C
farinosa2	.	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa3	.	A	.	C	T	.	.	.	.	G	.	.	.	.	**Y**	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa4	.	A	.	C	T	.	.	.	.	G	.	.	.	.	T	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa5	.	A	.	C	T	.	.	.	.	G	.	.	.	.	**Y**	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa6	.	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	**R**	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa7	.	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
farinosa8	.	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	-	.	-	-	A	C
farinosa9	.	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	-	.	-	-	A	C
halleri1	.	A	.	C	.	.	C	.	.	G	.	C	.	.	.	C	T	.	T	G	.	.	T	.	.	.	.	-	-	A	C
halleri2*	.	A	.	C	.	.	C	.	.	G	.	C	.	.	.	C	T	.	T	G	.	.	T	.	.	.	.	-	-	A	C
scotica1*	**R**	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
scotica2*	A	A	.	C	T	.	.	.	.	G	.	.	.	.	.	.	T	.	.	G	.	.	T	.	.	.	.	-	-	A	C
scandinavica1*	.	.	.	C	.	.	C	.	.	G	.	C	.	.	.	A	T	.	T	G	.	.	T	.	.	.	.	-	-	A	C
scandinavica2*	.	**S**	.	C	**Y**	.	**M**	.	.	G	.	**Y**	.	.	.	**R**	T	.	**Y**	G	.	.	T	.	.	.	.	-	-	A	C
stricta1*	**W**	**R**	.	C	**Y**	.	**M**	.	.	G	.	**Y**	.	.	.	**R**	T	.	**Y**	G	.	.	T	.	.	-	.	-	-	A	C
stricta2	**W**	**R**	.	C	**Y**	.	**M**	.	.	G	.	**Y**	.	.	.	**R**	T	.	**Y**	G	.	.	T	.	.	-	.	-	-	A	C

Appendix 2 (continued)

Accession	138	153	156	160	167	173	175	185	186	193	198	200	202	203	208	219	220	228	229	230	231	235	238	378	391	413	419	445	446	457	471
mistassinica1	C	C	C	C	G	C	C	A	T	G	A	C	C	G	T	C	G	C	T	C	G	C	A	T	C	G	T	C	C	C	G
mistassinica2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica3	.	.	.	.	.	.	.	.	.	.	.	.	**Y**	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
mistassinica4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
mistassinica5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica6	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica8	.	.	.	.	.	.	.	.	.	.	.	.	**Y**	**K**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica9*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
specuicola	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	C	.	.	.	T
alcalina*	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	T	.
magellanica1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
magellanica2*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
magellanica3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
stricta3*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
modesta	T	T	.	.	.	.	.	.	.	.	.	A	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
anvilensis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
borealis1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
borealis2*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
incana1*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
incana2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
incana3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
laurentiana1*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
laurentiana2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
laurentiana3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.
laurentiana4*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	**R**	C	.	.	.	.	.	.	.
farinosa1*	.	.	.	.	A	.	T	.	.	A	.	.	.	.	.	.	.	**Y**	**K**	.	.	G	.	C	.	.	.	.	.	.	.
farinosa2	.	.	.	.	A	.	T	.	.	A	.	.	.	.	.	.	.	T	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa3	.	.	.	.	A	**Y**	T	.	**W**	A	.	.	.	.	.	.	.	**Y**	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa4	.	.	.	.	A	.	T	.	A	A	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa5	.	.	.	.	A	**Y**	T	.	**W**	A	.	.	.	.	.	.	.	**Y**	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa6	.	.	.	.	A	.	T	.	.	A	.	.	.	.	.	.	.	T	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa7	.	.	.	.	A	T	T	.	.	A	.	.	.	.	.	.	.	T	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa8	.	.	.	.	A	.	T	.	.	A	.	.	.	.	.	.	.	**Y**	.	.	.	G	.	C	.	.	.	.	.	.	.
farinosa9	.	.	.	.	A	.	T	.	.	A	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.
halleri1	.	.	.	.	.	.	T	.	.	A	.	.	T	.	A	.	.	.	.	.	T	.	.	C	.	.	.	.	.	.	.
halleri2*	.	.	**Y**	.	.	.	T	.	.	A	.	.	T	.	A	.	.	.	.	.	T	.	.	C	.	**R**	.	.	.	.	.
scotica1*	.	.	.	**M**	A	.	T	.	.	A	**M**	.	.	.	.	T	.	.	.	.	.	**S**	.	C	.	.	.	.	**Y**	.	.
scotica2*	.	.	.	A	A	.	T	.	.	A	C	.	.	.	.	T	.	.	.	.	.	.	.	C	.	.	.	.	T	.	.
scandinavica1*	.	.	.	.	.	.	T	.	.	A	.	.	T	.	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
scandinavica2*	.	.	.	.	**R**	.	T	.	.	A	.	.	**Y**	.	**W**	.	.	.	.	.	.	**S**	.	**Y**	.	.	.	**M**	.	.	.
stricta1*	.	.	.	**M**	**R**	.	T	.	.	A	**M**	.	**Y**	.	**W**	**Y**	.	.	.	.	.	.	.	C	.	.	.	.	**Y**	.	.
stricta2	.	.	.	**M**	A	.	T	.	.	A	C	.	**Y**	.	**W**	**Y**	.	.	.	.	.	.	.	C	.	.	.	.	**Y**	.	.

Appendix 2 (continued)

Accession	489	498	499	500	504	514	517	522	532	536	537	554	571	574	581	583	588	592	594	595	606	617	621	623	624	632
mistassinica1	G	T	G	C	C	A	A	C	C	A	C	T	C	T	T	C	C	C	G	C	A	T	T	G	C	C
mistassinica2	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica3	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica4	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica5	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica6	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
mistassinica8	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
mistassinica9*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**R**	.	.	.	.	.	.	.
specuicola	C	.	.	.	.	.	.	.	.	G	.	.	.	C	A	.	.	.	.	.	.	C	.	.	.	.
alcalina*	.	.	T	.	.	.	C	.	.	G	.	.	.	C	.	.	.	.	.	.	.	C	.	A	.	.
magellanica1	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	A	.	.	.	.	.	C	.	.	.	.
magellanica2*	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	A	.	.	.	.	.	C	.	.	.	.
magellanica3	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	A	.	.	.	.	.	C	.	.	.	.
stricta3*	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
modesta	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	C	.	.	.	.
anvilensis	.	.	.	.	.	.	.	.	.	G	.	C	.	C	.	T	.	T	.	.	.	C	.	.	.	T
borealis1	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
borealis2*	.	.	.	.	.	.	.	.	.	G	.	C	**Y**	**Y**	.	.	.	.	.	.	.	**Y**	.	.	.	.
incana1*	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
incana2	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
incana3	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
laurentiana1*	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
laurentiana2	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
laurentiana3	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
laurentiana4*	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
farinosa1*	.	C	.	.	.	.	.	.	.	G	**Y**	C	.	.	.	.	**Y**	.	.	T	.	C	.	.	**M**	.
farinosa2	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	T	.	.	T	.	C	.	.	.	.
farinosa3	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	**Y**	.	.	T	.	C	.	.	.	.
farinosa4	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	.	.	.	T	.	C	.	.	.	.
farinosa5	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	**Y**	.	.	T	.	C	.	.	.	.
farinosa6	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	T	.	.	T	**R**	C	.	.	.	.
farinosa7	.	C	.	.	.	.	.	.	.	G	T	C	.	.	.	.	T	.	.	T	.	C	.	.	.	.
farinosa8	.	C	.	.	.	.	.	.	.	G	**Y**	C	.	.	.	.	**Y**	.	.	T	.	C	.	.	.	.
farinosa9	.	C	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	T	.	C	.	.	.	.
halleri1	.	C	.	.	.	T	.	T	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
halleri2*	.	C	.	.	.	T	.	T	**Y**	G	.	C	.	.	.	.	.	.	**R**	.	.	C	**W**	.	.	.
scotica1*	.	C	.	.	**Y**	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
scotica2*	.	C	.	.	T	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
scandinavica1*	.	C	.	T	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.
scandinavica2*	.	C	.	**Y**	.	.	.	.	.	G	.	C	.	.	.	.	.	.	**R**	.	.	C	.	.	.	.
stricta1*	.	C	.	**Y**	**Y**	.	.	.	.	G	.	C	.	.	.	.	.	.	.	.	.	C	.	.	.	.

Appendix 3. Polymorphic sites detected among clones of ITS amplicons: 25 additional polymorphic sites were detected that had not been recovered via direct sequencing. In those cases, polymorphisms usually consisted of a dominant nucleotide present in more than 75% of the sequenced clones, likely explaining the failure to identify the polymorphism in direct sequences of PCR products. Polymorphisms identified from direct sequences of ITS amplicons are highlighted in grey (IUPAC ambiguity codes: M=A or C, R=A or G, W=A or T, S=C or G, Y=C or T, K=G or T). Site numbering follows the ITS alignment. Sites identical to the first (i.e. direct) sequence (in bold) are represented by dots. Sites belonging to the 5.8S gene and separating ITS1 from ITS2 are italicised. Accession abbreviations as in Appendix 1.

	23	**27**	**31**	**33**	**39**	**40**	**42**
Accession	**5**	**9**	**3**	**6**	**3**	**5**	**5**
alcalina	**C**	**T**	**A**	**A**	**G**	**G**	**G**
clone1	.	.	.	.	A	A	.
clone2	.	C	.	.	.	.	.
clone3, 5, 7-13, 15-18, 20	.	.	.	.	.	.	.
clone4, 6	.	.	.	.	.	.	A
clone14	T	.	G	.	.	.	.
clone19	.	.	.	C	.	.	.

	443	**176**	**220**	**255**	**260**	**269**	**338**	**454**	**571**	**574**	**617**
Accesssion											
borealis2	**A**	**G**	**T**	**T**	**C**	**G**	**A**	**C**	**Y**	**Y**	**Y**
clone1, 5, 13	.	.	.	.	.	.	.	.	C	C	T
clone2-3, 7, 11	.	.	.	.	.	.	.	.	T	T	C
clone4	–	.	.	.	.	.	.	.	C	C	T
clone6	.	.	.	.	.	.	G	.	T	T	C
clone8	.	.	.	.	.	.	.	T	C	T	C
clone9-10	.	.	.	.	.	.	.	.	T	T	T
clone12	.	.	.	–	.	.	.	.	T	T	C
clone14	.	A	G	.	A	A	.	.	T	T	C
clone15-16	.	.	.	.	.	.	.	.	C	T	C

	68	**23**	**26**	**29**	**33**	**34**	**37**	**40**	**46**	**62**	**63**
Accession											
egaliksensis7	**A**	**T**	**C**	**C**	**T**	**C**	**C**	**G**	**T**	**T**	**C**
clone1-4, 6-7, 9, 11, 14-15, 17-18	.	.	.	.	.	.	.	.	.	.	.
clone5	.	C	.	.	C	.	.	.	.	.	.
clone8	.	.	A	.	.	.	.	.	.	.	.
clone10	.	.	.	.	.	.	.	.	.	.	G
clone12	.	.	.	.	.	.	T	.	.	C	.
clone13	.	.	.	.	.	A	.	.	.	.	.
clone16	G	.	.	.	.	.	.	A	.	.	.
clone19	.	.	.	G	.	.	.	.	C	.	.

Appendix 3 (continued)

Accession	1	38	84	108	109	123	139	149	152	159	166	186	203	228	229	252	301	441	457	463	537	548	588	624
farinosa1	**T**	**Y**	**C**	**R**	**W**	**G**	**A**	**A**	**A**	**A**	**C**	**T**	**G**	**Y**	**K**	**T**	**A**	**G**	**C**	**T**	**Y**	**G**	**Y**	**M**
clone1, 3, 8, 12	.	T	.	G	T	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	T	.	T	A
clone2	.	C	.	G	T	.	G	.	.	.	.	.	.	C	G	.	.	.	.	C	C	.	C	C
clone4	.	C	.	A	A	.	.	.	.	.	.	.	.	C	G	.	.	.	.	.	C	.	C	C
clone5	.	C	T	G	T	.	.	G	.	.	.	A	.	C	T	.	.	.	.	.	T	.	C	C
clone6	.	T	.	G	T	A	.	.	.	.	.	.	.	T	T	.	.	.	.	.	T	.	T	A
clone7, 10	.	C	.	A	A	.	.	.	G	.	.	.	.	C	G	.	T	.	.	.	C	.	C	C
clone9	.	C	.	A	A	.	.	.	.	.	T	.	.	C	G	.	.	.	.	.	C	.	T	C
clone11	.	C	.	A	A	.	.	.	.	.	.	.	.	C	G	.	.	A	.	.	C	.	C	C
clone13-14, 20-21, 24	.	C	.	A	A	.	.	.	.	.	.	.	.	C	G	.	.	.	.	.	C	.	C	C
clone15-16	.	C	.	A	A	.	.	.	.	.	.	C	.	T	T	.	.	.	.	.	T	.	T	A
clone17	.	C	.	A	A	.	.	.	.	.	.	.	.	C	G	.	.	.	T	.	T	.	T	A
clone18, 22	.	T	.	G	T	.	.	.	.	G	.	.	.	T	T	.	.	.	.	.	T	.	T	A
clone19	.	T	.	G	T	.	.	.	.	.	.	.	.	T	T	C	.	.	.	.	T	A	T	A
clone23	C	C	.	A	A	.	.	.	.	.	.	.	A	C	G	.	.	.	.	.	C	.	C	C

Accession	902	1259	1396	1560	1665	2000	2023	3101	4032	4041	4113	4242	4461	5222	5238	5322	5426	5549	6618	6221
halleri2	**C**	**G**	**A**	**Y**	**C**	**G**	**C**	**T**	**A**	**T**	**G**	**R**	**T**	**G**	**T**	**T**	**Y**	**G**	**R**	**C**	**W**
clone1, 16	.	.	.	C	.	.	.	.	.	C	.	A	.	.	.	.	C	.	A	.	T
clone2	.	.	T	C	.	.	.	.	.	.	.	A	.	.	.	.	T	T	G	T	T
clone3	.	–	.	C	.	.	.	.	.	.	.	G	.	.	.	.	C	.	G	.	A
clone4	.	.	.	C	.	.	.	.	.	.	–	G	.	.	.	.	C	.	G	.	T
clone5	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	C	C	.	A	.	T
clone6, 15	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	C	.	G	.	A
clone7	T	A	.	C	.	.	T	C	.	.	.	G	.	.	.	.	C	.	G	.	A
clone8	.	.	.	C	.	.	.	.	.	.	.	G	.	.	C	.	C	.	G	.	T
clone9	.	.	.	C	.	.	.	.	.	.	.	G	C	.	.	.	C	.	G	.	A
clone10, 19	.	.	.	T	.	.	.	.	.	.	.	A	.	.	.	.	T	T	G	.	T
clone11	.	.	.	T	.	T	.	.	.	.	.	A	.	.	.	.	T	T	G	.	T
clone12	.	.	.	C	.	.	.	.	G	.	.	G	.	.	.	.	C	.	G	.	A
clone13, 18	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	C	.	G	.	T
clone14	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	T	T	G	.	T
clone17	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	.	C	.	?	?	?
clone20	.	.	.	C	A	.	.	.	.	.	.	G	.	.	.	.	C	.	G	.	A
clone21	.	.	.	C	.	.	.	.	.	.	.	G	.	T	C	.	C	.	G	.	T

Accession	58	109	123	131	155	170	199	203	216	233	369	374	422	508	521	563
incana1	**R**	**Y**	**G**	**A**	**C**	**A**	**C**	**G**	**A**	**A**	**T**	**T**	**C**	**C**	**G**
clone1, 17	A	C	.	.	.	.	.	.	.	.	-	.	.	.	.
clone2	A	C	.	.	.	.	T	.	.	.	.	.	.	.	.
clone3	G	T	.	.	.	.	.	A	.	.	.	.	.	.	.
clone4, 8, 13-14	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.
clone5, 10	G	C	C	.	.	.	.	.	.	G	.	.	.	.	.
clone6	G	T	.	G	.	.	.	.	.	.	.	.	.	.	T
clone7	G	T	.	.	.	.	.	.	G	.	.	.	.	.	.
clone9	G	C	.	.	.	.	.	.	.	.	.	.	.	.	.
clone11	G	T	.	.	.	.	.	.	.	.	.	C	.	.	.
clone12	A	C	.	.	.	G	.	.	.	.	.	.	T	.	.
clone15, 18	A	C	.	.	T	.	.	.	.	.	.	.	.	.	.
clone16	A	C	.	.	.	.	.	.	.	.	.	.	.	T	.

Appendix 3 (continued)

Accession	42	121	126	136	150	151	163	193	202	218	220	232	233	250	265	266	288	418	519	553	558	619
laurentiana1	**A**	**G**	**C**	**C**	**C**	**A**	**C**	**G**	**C**	**T**	**T**	**G**	**A**	**T**	**C**	**T**	**T**	**A**	**T**	**T**	**C**	**A**
clone1, 9, 11, 16, 18-19, 21-22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone2	G	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
clone3	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
clone4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
clone5	.	.	–	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone6, 12	.	.	.	.	.	.	–	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
clone8, 17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
clone10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
clone13	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
clone14	.	.	.	.	.	.	.	A	A	.	G	T	G	A	A	.	C	.	.	C	.	.
clone15	.	.	.	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.

Accession	17	154	162	172	176	227	235	238	288	297	550	625
laurentiana4	**T**	**A**	**C**	**C**	**G**	**G**	**C**	**R**	**T**	**T**	**T**	**T**
clone1-2	C	.	.	T	.	.	.	G	.	.	.	.
clone3, 5, 9, 17	.	.	.	.	.	.	.	G	.	.	.	.
clone4	.	.	T	.	.	.	.	G	.	.	.	.
clone6	.	.	.	.	.	.	.	G	.	.	.	?
clone7	.	.	.	.	.	T	.	A	.	.	.	.
clone8	.	.	T	.	.	.	.	G	.	.	.	C
clone10, 18, 20	.	G	.	.	.	.	.	G	.	.	.	.
clone11	.	.	.	.	.	.	.	G	.	C	.	.
clone13	.	.	.	.	.	.	G	A	.	.	.	.
clone14	.	.	.	.	.	.	.	A	C	.	.	.
clone15	.	.	.	.	.	.	G	A	.	.	.	.
clone16	.	.	.	.	.	.	.	G	.	.	C	C
clone19	.	.	.	.	A	.	.	G	.	.	.	.

Accession	54	170	287	341	496	557
magellanica2	**A**	**A**	**C**	**C**	**T**	**A**
clone1	G	.	.	G	.	.
clone2	.	.	T	.	.	.
clone3	.	.	.	.	C	G
clone4	G	.	.	.	.	.
clone5	.	G	.	.	.	.
clone6-9	.	.	.	.	.	.

Accession	89	804	892	288	300	331	332	355	562	594	616
mistassinica9	**T**	**T**	**Y**	**C**	**T**	**A**	**A**	**T**	**T**	**R**	**A**
clone1, 4, 9	.	.	C	.	.	.	.	.	.	A	.
clone2	.	.	C	.	.	G	.	.	.	A	.
clone3	.	.	C	.	.	.	.	C	.	A	.
clone5	.	C	C	.	.	.	.	.	.	A	G
clone6	C	.	C	.	C	.	G	.	C	G	.
clone7	.	.	T	T	.	.	.	.	.	G	.
clone8	.	.	C	.	.	.	.	A	.	A	.

Appendix 3 (continued)

Accession	12	43	46	80	102	126	148	156	179	202	239	246	262	284	290	311	327	344	358	375	385	403	414
nutans7	**G**	**A**	**R**	**T**	**G**	**T**	**C**	**Y**	**G**	**C**	**A**	**T**	**A**	**G**	**G**	**Y**	**R**	**G**	**G**	**Y**	**T**	**C**	**G**
clone1	A	.	G	.	A	C	.	C	.	T	T	.	.	A	.	T	G	.	.	T	.	.	.
clone2	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	C	G	.	.	T	.	T	.
clone3	A	.	G	.	.	C	.	C	T	T	T	.	T	.	.	T	A	.	.	C	C	T	.
clone4	A	.	G	C	.	C	.	C	.	T	T	.	.	.	.	T	A	.	.	C	C	T	T
clone5, 7	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	C	G	.	.	T	.	.	.
clone6	.	.	G	.	.	.	.	C	.	.	.	.	.	.	.	C	G	.	A	T	.	.	.
clone8	A	–	G	.	.	C	T	C	.	T	T	C	.	.	A	T	A	A	.	C	C	T	.

Accession	81	112	133	334	343	411	516	558	717	809	923	1024	1216	1328	1434	1491	1563	1593	1615
scandinavica1	**C**	**T**	**G**	**C**	**A**	**C**	**T**	**C**	**A**	**A**	**A**	**G**	**A**	**T**	**A**	**A**	**G**	**T**	**A**	**C**
clone1, 5, 7, 9, 14, 17-18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
clone3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.	.	.	.
clone4, 8	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.
clone6	.	A	A	T	.	A	–	.	.	.	.	.	.	.	.	.	.	.	.	.
clone10	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.
clone11	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	T	.	.	.
clone12	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
clone13	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	T
clone15	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	.	.	.	.
clone16	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.

Accession	68	12	22	33	38	43	48	54	61	62	85	105	107	115	116	120	128	135	139	142	147	148	149	150	154	155	159	164	173	180
scandinvica2	**A**	**C**	**S**	**Y**	**A**	**M**	**G**	**A**	**Y**	**G**	**R**	**Y**	**R**	**A**	**A**	**Y**	**W**	**S**	**T**	**C**	**Y**	**M**	**G**	**Y**	**A**	**A**	**A**	**R**	**T**	**T**
clone1	.	T	G	C	.	C	.	.	C	.	A	T	G	.	.	T	A	C	.	.	C	C	.	T	.	.	G	G	.	.
clone2	.	.	G	C	.	C	.	.	C	.	A	T	G	.	.	T	A	C	.	.	C	C	A	T	.	.	.	G	.	.
clone3	.	.	C	T	.	A	.	.	T	.	G	C	A	.	.	T	A	C	.	.	C	C	.	C	T	.	.	G	.	.
clone4	G	.	C	T	.	A	.	.	T	T	G	C	A	.	.	C	T	G	.	.	T	A	.	C	.	.	.	A	.	.
clone5, 11, 13-14, 18, 20, 22-23	.	.	C	T	.	A	.	.	T	.	G	C	A	.	.	C	T	G	.	.	T	A	.	C	.	.	.	A	.	.
clone6	.	.	C	T	.	A	.	.	T	.	G	C	A	.	.	C	T	G	.	.	T	A	.	C	.	.	.	A	.	C
clone7	.	.	C	T	.	A	.	.	T	.	G	C	A	.	.	C	T	G	.	.	T	A	.	C	.	.	.	G	.	.
clone8	.	.	G	C	G	C	.	.	C	.	A	T	G	.	.	T	A	C	.	.	C	C	.	T	.	.	.	G	.	.
clone9	.	.	G	C	.	C	.	.	C	.	A	T	G	.	.	T	A	C	.	.	C	C	.	T	.	.	.	G	.	.
clone10	.	.	C	T	.	A	.	G	T	.	G	C	A	.	.	T	A	C	.	.	T	A	.	C	.	.	.	A	.	.
clone12	.	.	C	T	.	C	.	.	C	.	A	T	G	.	.	C	T	C	C	.	C	C	.	T	.	.	.	G	.	.
clone15	.	.	C	T	.	A	A	.	T	.	G	C	A	.	.	C	T	G	.	.	T	A	.	C	.	.	.	A	.	.
clone16	.	.	G	C	.	C	.	.	C	.	A	T	G	.	.	T	T	C	.	.	C	C	.	T	.	.	.	G	.	.
clone17	.	T	G	C	.	C	.	.	C	.	A	T	G	G	.	T	A	C	.	.	C	C	.	T	.	.	.	A	.	.
clone19	.	.	G	C	.	C	.	.	C	.	A	T	G	.	G	T	A	C	.	.	C	C	.	T	.	.	.	A	.	.
clone21	.	.	G	C	.	C	.	.	C	.	A	T	G	.	.	T	A	G	.	T	T	A	.	C	.	T	.	A	C	.

Appendix 3 (continued)

Accession	11	12	32	43	45	61	73	85	90	105	115	123	125	160	164	167	198	208	219	231	235	237	308	320	353	375	378	406	445	463
scotica1	**R**	A	T	A	C	T	C	G	C	C	T	G	G	**M**	G	A	**M**	T	T	G	**S**	T	G	A	C	G	C	C	C	**Y**	T
clone1	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone2	T	C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	C	.	G	.	A	.	.	.	T	T	A	C	.
clone3	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone4	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone5	T	.	C	C	A	C	.	C	T	T	.	.	A	C	.	G	A	A	C	T	C	.	.	.	.	.	.	.	.	C	C
clone6	A	.	.	.	.	.	.	.	.	.	.	.	.	A	T	.	C	.	.	.	C	.	.	.	T	.	.	.	.	T	.
clone7	G	.	.	.	.	.	A	.	.	.	C	.	.	A	.	.	C	.	.	.	C	A	.	.	.	C	.	.	.	T	.
clone8, 10, 15	A	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone9	A	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone11	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	G	.	.	.	.	.	T	.
clone12	T	.	C	C	.	C	.	C	T	T	.	.	A	C	.	G	A	A	C	T	C	.	.	.	.	.	.	.	.	C	.
clone13	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	T	.	A	C	.
clone14	G	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	T	.
clone16	T	C	.	.	.	.	.	.	.	.	.	T	-	C	.	.	A	.	C	.	G	.	.	.	.	C	T	.	A	C	.

Accession	470	472	483	489	504	510	544	555	567	594
scotica1	G	A	C	C	**Y**	A	A	A	G	G
clone1	.	.	.	.	T	.	.	.	.	.
clone2	.	.	.	.	C	.	.	.	.	A
clone3	.	.	.	.	C	C	.	.	.	A
clone4	.	.	A	.	T	.	.	.	A	.
clone5	.	.	.	.	C	.	T	.	.	.
clone6	.	.	.	T	T	.	.	.	.	.
clone7	.	.	.	.	T	.	.	.	.	.
clone8, 10, 15	.	.	.	.	T	.	.	.	.	.
clone9	.	.	.	.	C	.	T	.	.	.
clone11	A	.	.	.	T	.	.	G	.	.
clone12	.	.	.	.	C	.	T	.	.	.
clone13	.	.	.	.	T	.	.	.	.	.
clone14	.	T	.	.	T	.	.	.	.	.
clone16	.	.	.	.	C	.	.	.	.	A

Accession	131	136	178	187	175	203	201	225	243	275	293	337	404	513
scotica2	A	T	G	C	T	G	G	A	G	T	G	C	G	A
clone1	.	C	.	.	.	.	.	.	.	.	.	.	.	.
clone2	.	.	.	.	.	.	A	.	.	.	.	.	.	.
clone3-4, 7, 12, 16-17	.	.	.	.	.	.	.	.	.	.	.	.	.	.
clone5	.	.	.	.	.	.	.	G	.	C	.	.	.	.
clone6	.	.	.	.	.	A	.	.	.	.	.	.	.	.
clone8	G	.	.	.	.	.	.	.	A	.	C	.	.	G
clone9	.	.	.	.	.	.	.	.	.	.	.	.	T	.
clone10	.	.	A	.	.	.	.	.	.	.	.	.	.	.
clone11, 14	G	.	.	.	.	.	.	.	.	.	.	.	.	.
clone13	.	.	.	.	C	.	.	.	.	.	.	T	.	.
clone15	.	.	.	.	.	.	.	.	.	.	.	.	T	.
clone18	.	.	.	T	.	.	.	.	.	.	.	.	.	.

Appendix 3 (continued)

Accession	11	12	23	34	61	73	79	85	105	112	134	145	156	160	167	174	189	192	200	207	228	239	243	253	263	279	290	306	314	326	335	341	354	360	374	389
stricta1	**W**	**R**	**Y**	**M**	**Y**	C	G	**R**	**Y**	**–**	C	C	**M**	**R**	A	**M**	**Y**	C	**W**	**Y**	A	C	T	C	A	G	**Y**	**Y**	**Y**	C	T	A
clone1, 3	T	G	C	C	C	.	.	A	T	G	.	.	C	G	.	A	G	.	A	C	G	.	.	.	.	.	C	T	C	.	.	.
clone2	A	A	T	A	T	.	.	G	C	**–**	T	.	A	A	.	C	C	.	T	T	.	.	.	.	.	.	T	C	T	.	.	.
clone4, 8	A	A	T	A	T	.	.	G	C	**–**	.	T	A	A	.	C	C	.	T	T	.	.	.	.	.	.	T	C	T	A	.	.
clone5, 11	T	A	T	A	T	.	A	G	C	**–**	.	.	A	A	G	A	C	T	T	C	.	G	.	A	.	.	T	C	T	.	.	.
clone6	T	G	C	C	C	.	.	A	T	G	.	.	C	G	.	A	T	.	A	C	.	.	.	.	.	.	C	T	C	.	C	.
clone7	A	A	T	A	T	T	T	G	C	**–**	.	.	A	A	.	C	C	.	T	T	.	.	.	.	T	.	T	C	T	.	.	.
clone9, 12	T	G	C	C	C	.	.	A	T	G	.	.	C	G	.	A	T	.	A	C	.	.	.	.	.	.	C	T	C	.	.	.
clone10	A	A	T	A	T	.	.	G	C	**–**	.	.	A	A	.	C	C	.	T	T	.	.	.	.	.	.	T	C	T	.	.	.
clone13	T	G	C	C	C	.	.	A	T	G	.	.	C	G	.	A	T	.	A	C	.	.	C	.	.	.	T	C	T	.	.	.
clone14	T	G	C	C	C	.	.	A	T	G	.	.	C	G	.	A	T	.	A	C	.	.	.	.	.	T	C	T	C	.	.	G

Accession	63	119	133	152	160	175	180	194	200	219	226	237
stricta3	**G**	**C**	**G**	**G**	**A**	**C**	**T**	**C**	**T**	**C**	**A**	**T**
clone1	A	.	.	.	.	.	.	.	.	T	.	.
clone2-4, 7, 13-14, 16-18	.	.	.	.	.	.	.	.	.	.	.	.
clone5	.	T	.	.	.	.	.	.	.	.	.	?
clone6, 10, 12	.	.	.	.	.	.	.	.	.	.	.	?
clone8	.	.	.	A	.	.	.	.	.	.	.	.
clone9	.	.	.	.	.	T	.	T	C	.	.	.
clone11, 19	.	.	.	.	.	.	C	.	.	.	G	?
clone15	.	.	.	.	.	.	.	.	.	.	.	C
clone20	.	.	A	.	G	.	.	.	.	.	.	.

Chapter 4

Genomic origin and organisation of the allopolyploid *Primula egaliksensis* investigated by *in situ* hybridisation

with C. Baroux, U. Grossniklaus, and E. Conti

Annals of Botany (2008) doi:10.1093/aob/mcn026

Abstract

Earlier studies suggested that the tetraploid *Primula egaliksensis* (2n=40) originated from hybridisation between the diploids *P. mistassinica* (2n=18) and *P. nutans* (2n=22), which were hypothesised to be the maternal and paternal parent, respectively. The present paper is aimed at verifying the hybrid nature of *P. egaliksensis* with cytogenetic tools, and investigating the extent to which the parental genomes underwent genomic reorganisation. Genomic *in situ* hybridisation (GISH) and fluorescent *in situ* hybridisation (FISH) with ribosomal DNA (rDNA) probes, together with sequencing of the internal transcribed spacer (ITS) region of the rDNA, were used to identify the origin of *P. egaliksensis* and explore its genomic organisation, particularly at rDNA loci. GISH showed that *P. egaliksensis* inherited all chromosomes from *P. mistassinica* and *P. nutans* and did not reveal major intergenomic rearrangements between the parental genomes (e.g. interchromosomal translocations). However, karyological comparisons and FISH experiments suggested small-scale rearrangements, particularly at rDNA sites. *Primula egaliksensis* lacked the ITS-bearing heterochromatic knobs characteristic of the maternal parent *P. mistassinica* and maintained only the rDNA loci of *P. nutans*. These results corroborated sequence data indicating that most ITS sequences of *P. egaliksensis* were of the paternal repeat type. The lack of major rearrangements may be a consequence of the considerable genetic divergence between the putative parents, while the rapid elimination of the ITS repeats from the maternal progenitor may be explained by the subterminal location of ITS loci or a potential role of nucleolar dominance in chromosome stabilisation. These small-scale rearrangements may be indicative of genome diploidisation, but further investigations are needed to confirm this assumption.

Introduction

Polyploidisation has played a major role in plant speciation (Soltis and Soltis, 1993; Bretagnolle *et al.*, 1998; Otto and Whitton, 2000; Levin, 2002; Soltis *et al.*, 2003). At least 70% of all angiosperms are thought to have polyploid origins (reviewed by Soltis, 2005) including species of small genome size and chromosome number such as *Arabidopsis thaliana* (Henry *et al.*, 2006). Two types of polyploids are commonly recognised according to the degree of homology among coexisting genomes. Autopolyploids, which arise within a single species, contain more than two sets of homologous chromosomes in their nuclear genome, while allopolyploids contain more than two sets of homeologous chromosomes that diverged from each other prior to interspecific hybridisation (Ramsey and Schemske, 1998).

Polyploids generally undergo rapid genome restructuring following their formation (Wendel, 2000; Adams and Wendel, 2005; Chen and Ni, 2006) as demonstrated, for instance, by the studies on synthetic and natural allopolyploids of *Nicotiana* (Skalická *et al.*, 2003; Kovarík *et al.*, 2004; Lim *et al.*, 2004a; Skalická *et al.*, 2005), wheat (Levy and Feldman, 2004) and *Arabidopsis* (Pontes *et al.*, 2004). Genomic changes range from intergenomic chromosome translocations to interlocus recombinations (e.g. unequal crossing-over and gene conversion), and can be accompanied by epigenetic modifications (e.g. cytosine methylation) at specific loci leading to altered gene expression patterns (Wendel, 2000; Liu and Wendel, 2003; Wang *et al.*, 2004; Chen and Ni, 2006). Altogether these changes may constitute a response to the cellular stress imposed by the novel hybrid chromosome composition and/or doubling of the gene dosage (i.e. genomic shock; McClintock, 1984; Madlung and Comai, 2004) and result in the gradual diploidisation of polyploid taxa (Leitch and Bennett, 2004; Ma and Gustafson, 2005). Yet, there are also examples of polyploids which have undergone only few changes in overall genome structure since their formation, e.g. in *Gossypium* (Liu *et al.*, 2001) and *Spartina* (Baumel *et al.*, 2002; Ainouche *et al.*, 2004).

The varied response of plant genomes to polyploidisation is reflected in the nature and arrangement of ribosomal DNA (rDNA) repeats of polyploids. Ribosomal genes occur in hundreds or thousands of copies that are tandemly repeated over one or more loci (Long and Dawid, 1980). Some hybrids show a strict conservation in number, but not necessarily in size, of both parental rDNA loci (e.g. in *Nicotiana*; Matyásek *et al.*, 2003; Kovarík *et al.*, 2004), while others display gain (e.g. in *Gossypium*; Hanson *et al.*, 1996) or loss (e.g. in *Arabidopsis*; Pontes *et al.*, 2004) of rDNA loci. At the sequence level, similar contrasting scenarios of rDNA loci evolution were observed as sequences from putative allopolyploids were found to be either additive for differing parental nucleotides (e.g. Campbell *et al.*, 1997), or homogenised towards one parental repeat type (e.g.

Wendel *et al.*, 1995; Kovarík *et al.*, 2004), due to processes of intra- and interlocus recombination (referred to as concerted evolution; Elder and Turner, 1995) or loss of rDNA loci.

The genus *Primula* (totalising ca. 430 species; Richards, 2002) constitutes an ideal model system to investigate the effects of hybridisation and polyploidisation on plant evolution. On the one hand, primulas have been hybridised for gardening purposes since the 16th century and several sections are known for their high levels of interspecific hybridisation in the wild (e.g. sects. *Auricula* and *Primula*; Richards, 2002). On the other hand, the ploidy level varies from diploid to 14-ploid, with some sections being entirely polyploid (e.g. sects. *Auricula* and *Parryi*; Richards, 2002). Interestingly, morphological, caryological, distributional and phylogenetic data suggest that most polyploids belonging to sect. *Aleuritia* are of allopolyploid origin (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Guggisberg *et al*., unpubl. res.).

The tetraploid *Primula egaliksensis* (2n=40) is widely distributed across North-America and is taxonomically ascribed to sect. *Armerina* owing to its entire petiolate leaves, narrow elongated capsules and the absence of farina (Kelso, 1991; Richards, 2002). Karyological and morphological evidence suggest, however, that it may be an intersectional allopolyploid involving one species (out of 27) of sect. *Aleuritia* and one species (out of 14) of sect. *Armerina* (Kelso, 1991, 1992). This hypothesis relies primarily on the additivity of chromosome numbers, the intermediacy of gland types, pollen sizes, colpi numbers and exine reticulation patterns (Kelso, 1991, 1992). Within sect. *Aleuritia*, four diploid species occur in North-America (*P. alcalina*, *P. anvilensis*, *P. mistassinica* and *P. specuicola*), but chloroplast DNA (cpDNA) based phylogenies suggest that *P. egaliksensis* is more closely related to *P. mistassinica* (2n=18) than to any other species of *Primula* (Fig. 1; Mast *et al.*, 2001; Guggisberg *et al.*, 2006; Mast *et al.*, 2006; Guggisberg *et al.*, unpubl. res.). Within sect. *Armerina*, *P. nutans* (2n=22) is the only diploid that occurs in North-America and phylogenetic evidence based on rDNA sequences indicate a common origin for *P. nutans* and *P. egaliksensis* (Fig. 1; Guggisberg *et al.*, unpubl. res.). Hence, *P. mistassinica* and *P. nutans* are the most likely progenitors of *P. egaliksensis*, the former probably acting as the maternal parent. Distributional data further suggest that *P. egaliksensis* originated in North-America, because it is rarely found outside of this area and the current ranges of its putative parents only overlap in north-western Canada. Finally, the diversity of cpDNA haplotypes recovered by sequencing of multiple accessions of *P. egaliksensis* advocates for a recurrent origin of this taxon (Guggisberg *et al.*, 2006).

The numerous studies on synthetic and natural polyploids have shown that polyploid genomes are dynamic in their response to polyploidisation, underlining the need for additional case studies looking at genomic, genetic and epigenetic changes following chromosome doubling. Cytogenetic investigations are powerful in verifying the origin of polyploid taxa and may provide preliminary

insights into the extent of genomic rearrangements. The present paper intends to (i) confirm the hypothesised allopolyploid origin of *P. egaliksensis*, (ii) investigate the genomic organisation of this tetraploid by genomic *in situ* hybridisation (GISH), and (iii) understand the evolutionary history of the 45S rDNA loci – consisting of the 18S, 5.8S and 26S ribosomal genes and the internal transcribed spacers (ITS) – using fluorescent *in situ* hybridisation (FISH) and ITS sequencing.

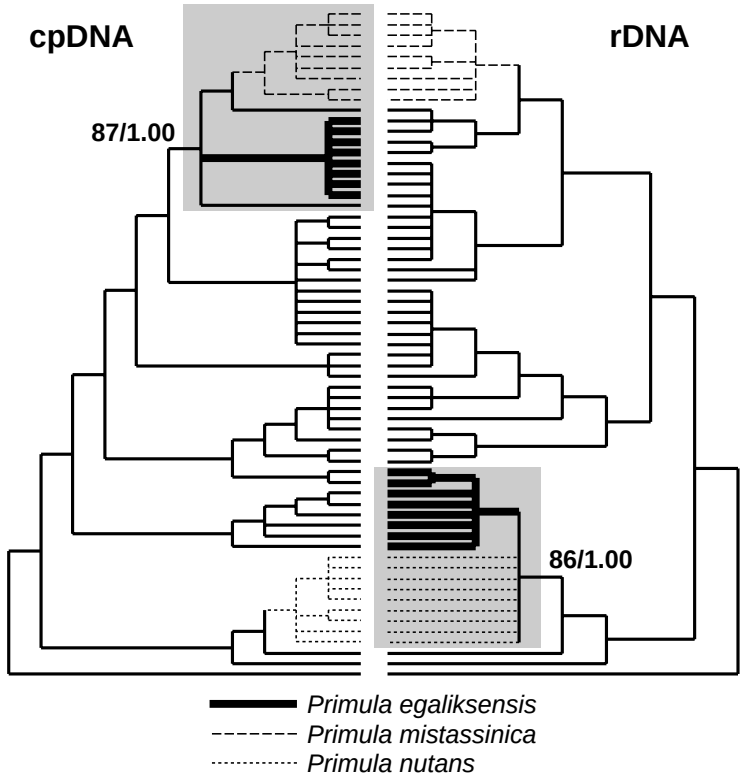


Figure 1. Phylogenetic placement of *Primula egaliksensis* in relation to its putative parents, i.e. *P. mistassinica* and *P. nutans*, in the cpDNA- and rDNA-based phylogenies inferred by Bayesian analysis. Clades of interest are highlighted in grey and followed by bootstrap support values and posterior probabilities, separated by a slash. Redrawn from Guggisberg *et al.* (unpubl. res.).

Materials and Methods

Plant material

The plant material used in this study is listed in Table 1. We used different accessions for DNA extractions and chromosome spreads, because none of the sampled populations contained both the leaf and seedling tissues necessary for the two sets of analyses.

DNA extraction, PCR amplification, cloning and sequencing

Total genomic DNA (gDNA) was extracted using the DNeasy Plant Mini Kit (Qiagen, Switzerland). The two ITS and the 5.8S ribosomal gene of the rDNA cistron (hereafter called the ITS region, covering 638 bp) were amplified with ITS.LEU and ITS4 primers (Baum *et al.*, 1998) using 2 mM $MgCl_2$, 200 µM dNTPs and 0.2 µM of each primer in a standard PCR buffer using an annealing temperature of 52°C (with a ramp speed of 1°C/s). PCR products destined for sequencing

were purified (GFX PCR DNA and Gel Band Purification Kit, Biosciences Amersham, Switzerland) and cloned into pCR®II-TOPO® (TOPO TA Cloning® kit, Invitrogen, Switzerland) before sequencing (ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit, ABI Prism 3100 automated sequencer, Applied Biosystems, USA). Because preferential amplification of one sequence variant may lead to its over-representation in the final reaction mixture (cf. PCR drift and PCR selection; Wagner *et al.*, 1994), three PCR reactions of each sample were pooled for use in ligations (Mason-Gamer, 2004).

Table 1. Plant material of *Primula* used in this study. Voucher specimens were deposited at the herbarium of Z, Zurich, Switzerland.

Taxon	Tissue type	Provenance	Herbarium voucher
P. egaliksensis	leaf	Alaska/USA	Guggisberg & Mansion 100703-1b[1]
	leaf	Newfoundland/Canada	Guggisberg & Mansion 240604-6b[2]
	root tip	Alaska/USA	Guggisberg & Mansion 060703-1
	root tip	Alaska/USA	Guggisberg & Mansion 150703-1
P. mistassinica	leaf	Wisconsin/USA	Anderson 210503-3
	root tip	Alberta/Canada	Eveleigh 300603-3
P. nutans	leaf	Alaska/USA	Guggisberg & Mansion 240603-2
	root tip	Yukon Territory/Canada	Guggisberg & Mansion 250603-2
	root tip	Yukon Territory/Canada	Bennett, Line & Hett-Seccombe 030604-1

[1]Abbreviated '*P. egaliksensis* 1' in Table 2. [2]Abbreviated '*P. egaliksensis* 2' in Table 2.

Sequencher 4.2 (Gene Codes Corp., USA) was used to check the quality of the electropherograms and compile the contiguous sequences for each PCR product. The starting and ending points of each sequence were determined by comparison with the partial rDNA sequence of *Rhododendron kanehirai* (GenBank AF172290) and deposited in GenBank (accession no. DQ993783-DQ993802, EU095369-EU095392). Cloned ITS sequences of *P. egaliksensis* were finally aligned with 18 sequences of *P. mistassinica* (GenBank DQ993741-DQ993749) and *P. nutans* (GenBank DQ993763-DQ993771) using Se-Al v2.0a9 (available at http://tree.bio.ed.ac.uk/), in order to determine the direction of concerted evolution at sites distinguishing the putative parents. The alignment was searched for variable positions using PAUP* 4.0.10b (Swofford, 1999).

Chromosome preparation and in situ *hybridisation*

Seeds were imbibed in water, stratified six weeks at 4°C, and germinated on Murashige and Skoog basal medium in a constantly illuminated growth chamber at 20°C. Two week old seedlings were treated with auxin (2.5 µM naphthalene acetic acid in medium) for one week to induce the formation of root meristems. Mitotic cells were stopped at the metaphase stage upon exposure to colchicine (0.05% in liquid medium) for two hours. Circa 1 cm root tip samples were excised and fixed in ice-cold ethanol:acetic acid (3:1, v/v) and kept at –20°C until use. Spread nuclei were prepared essentially as described by Lysak *et al.* (2006), but with 50% acetic acid.

The preparations were stained with 1 µg/mL 4,6-diamino-2-phenylindole (DAPI) in Vectashield antifade (Vector Laboratories, Canada) and screened for the quality of metaphase, prophase and interphase nuclei under epifluorescence microscopy (Axioplan, Zeiss, Germany and DM6000, Leica, Germany). The best preparations were rinsed in PBS (10 mM sodium phosphate, pH 7.0, 143 mM NaCl), dehydrated in ethanol series (70%, 90%, 100%, 2 min each), and air dried before FISH or GISH analyses.

For FISH and GISH experiments, slide and probe preparation, hybridisation and detection were done as described by Lysak *et al.* (2006) with the following minor modifications: the RNase treatment was done 30 min instead of 60 min; chromosomes were treated with 25 µg/mL pepsin in 0.01 M HCl for 2 min at 37 °C before hybridisation; the probes were pre-denatured 15 min at 75°C and cooled on ice for 5 min before hybridisation onto the chromosomes; post-hybridisation washes were done at 45°C for 10 min each, followed by rinses at room temperature; for GISH, the chromosomes were denatured in 70% formamide in 2X SSC for 2 min at 80°C, fixed in 70% ethanol at –20°C for 5 min and dehydrated in 90% and 100% ethanol at room temperature. The hybridisation mixture consisted of 50-100 ng of labelled probes, 50% deionised formamide, 10% dextran sulphate, 2X SSC, and 50 µg sheared salmon sperm DNA. Pre-denaturation of the genomic probe mix was found to be essential to eliminate cross-hybridisation at heterochromatic regions, possibly due to quick re-annealing of conserved repeat sequences in the probe mixture before target chromosomes were reached.

The 45S rDNA probe was prepared from the plasmid pTa71 (a gift from P. Franz, University of Amsterdam) containing the 18S-25S ribosomal genes isolated from *Triticum aestivum* L. (Barker *et al.*, 1988), and labelled with digoxigenin (Nick Translation Mix, Roche, Switzerland). The ITS probes of *P. egaliksensis* (ITS-PE), *P. mistassinica* (ITS-PM) and *P. nutans* (ITS-PN) were prepared from PCR products as described above, and labelled with biotin (Nick Translation Mix, Roche, Switzerland). ITS probes from *P. mistassinica* (ITS-PM) and *P. nutans* (ITS-PN) were used on their respective genomes, but also on nuclei of *P. egaliksensis*. For GISH, gDNA of *P. mistassinica* (gDNA-PM) and *P. nutans* (gDNA-PN) was extracted as described above, and sheared to ca. 1000 bp by sonication for 10 s (UW2200, Bandelin Electronic, Germany) before labelling with either digoxigenin- or biotin-labelled nucleotides (Nick Translation Mix, Roche, Switzerland). Digoxigenin-labelled probes were detected using a Mouse-anti-Digoxigenin (Roche, Switzerland) and a Goat-anti-Mouse antibody coupled to Alexa 488 (Molecular Probes, Invitrogen, Switzerland) as primary and secondary antibody, respectively (Lysak *et al.*, 2006). Biotin-labelled probes were detected with Avidin and Goat-anti-Avidin coupled to Texas Red as described by Lysak *et al.*

(2006). All images were processed using Adobe Photoshop® (Adobe Systems, Switzerland) and treated for colour contrast and brightness.

Results

GISH demonstrates the hybrid origin of P. egaliksensis

The investigation of root tip prophase and metaphase cells of *P. mistassinica*, *P. nutans* and *P. egaliksensis* confirmed previous reports of chromosome numbers. A diploid chromosome number was counted in both *P. mistassinica* (2n=18; Fig. 2a) and *P. nutans* (2n=22; Fig. 2b), and a tetraploid chromosome number was counted in *P. egaliksensis* (2n=40; Fig. 2c). In addition, most chromosomes of *P. mistassinica* were characterised by heterochromatic knobs (i.e. highly condensed chromatin) corresponding to brightly stained chromosomal regions (cf. arrowhead and inset in Figure 2a) that were distinct from pericentromeric regions. Noticeably, these genomic structures were absent in both *P. nutans* and *P. egaliksensis*.

In order to assess whether *P. egaliksensis* inherited the genome complements from both *P. mistassinica* and *P. nutans*, GISH was carried out. Control experiments using labelled gDNA of *P. nutans* to hybridise *P. mistassinica* nuclei, and vice versa, showed very little or no cross-hybridisation with our experimental conditions (see Materials and Methods; data not shown). Hence, the genomes of the putative parents were differentially labelled and hybridised to *P. egaliksensis* spread nuclei of root tips. The results unambiguously demonstrate the hybrid origin of *P. egaliksensis*, for interphase nuclei as well as metaphase chromosomes of *P. egaliksensis* clearly revealed the presence of both *P. nutans* (red) and *P. mistassinica* (green) genomes (Figs. 2d-j).

Parental chromosomes of P. egaliksensis *remain globally preserved*

GISH on *P. egaliksensis* metaphase plates allowed to count 18 chromosomes hybridised with gDNA of *P. mistassinica* and 22 chromosomes hybridised with gDNA of *P. nutans* as expected if all parental chromosomes were inherited by the allopolyploid (Figs. 2e-j). As evidenced by the relative homogenous hybridisation pattern on each chromosome, no chromosomes of *P. egaliksensis* presented hybridisation signals from both parents, suggesting that no major intergenomic rearrangements occurred between the parental genomes (e.g. interchromosomal translocations).

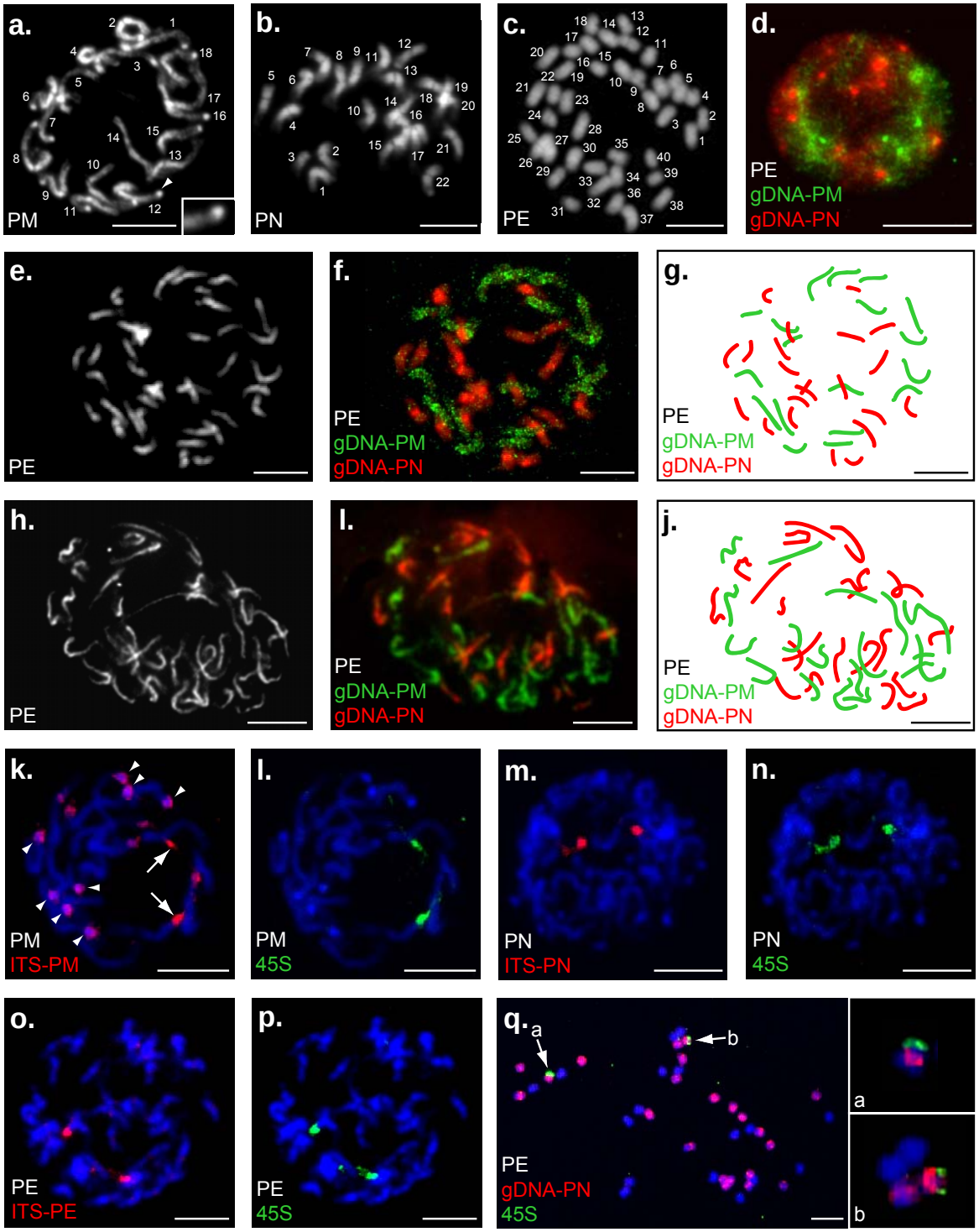


Figure 2. (a-c) DAPI stained root tip prophases/metaphases of (a) *Primula mistassinica* showing 2n=18 chromosomes, (b) *P. nutans* showing 2n=22 chromosomes, and (c) *P. egaliksensis* showing 2n=40 chromosomes; the arrowhead in (a) points to the heterochromatic knob shown in the inset. (d) GISH to interphase nucleus of *P. egaliksensis* using gDNA from *P. mistassinica* (green) and *P. nutans* (red). (e-j) Root tip prophases/metaphases of *P. egaliksensis* stained with DAPI (e, h) and resulting GISH signal (f-g, i-j) following *in situ* hybridisation of gDNA from *P. mistassinica* (green) and *P. nutans* (red). (k-p) FISH localisation of ITS (red) and 45S rDNA (green) loci on root tip prophases of (k-l) *P. mistassinica*, (m-n) *P. nutans*, and (o-p) *P. egaliksensis* counterstained with DAPI; in (k), arrows point to cross-hybridisation of ITS and 45S rDNA probes, and arrowheads highlight heterochromatic knobs. (q) Root-tip metaphase of *P. egaliksensis* counterstained with DAPI after GISH with gDNA of *P. nutans* (red) and FISH with 45S rDNA probes (green); arrows (q_{a-b}) point to the two 45S rDNA bearing chromosomes shown in the insets. Scale bar = 10 µm. PE, *P. egaliksensis*; PM, *P. mistassinica*; PN, *P. nutans*.

ITS sequences of P. egaliksensis *are mainly homogenised towards the* P. nutans *repeat type*

Sequencing of 44 ITS clones from two accessions of *P. egaliksensis* revealed 22 different ITS sequences. These clones varied at 21 of the 27 positions distinguishing the putative progenitors of *P. egaliksensis*, i.e. *P. mistassinica* and *P. nutans* (Table 2). Thirty-six clones (82%; no. 1-13/15/17-18/20-21 of *P. egaliksensis* 1 and no. 1-15/17-19 of *P. egaliksensis* 2) were completely homogenised towards the ITS repeat type found in *P. nutans*. Of the remaining eight clones, six (no. 14/16/19/22 of *P. egaliksensis* 1 and no. 16/20 of *P. egaliksensis* 2) were mainly, but not completely, homogenised towards the ITS repeat type found in *P. nutans*, while two (no. 21-22 of *P. egaliksensis* 2) were mainly homogenised towards the ITS repeat type found in *P. mistassinica*. Hence, only 5% of the clones presented a nucleotide composition characteristic of *P. mistassinica* ITS sequences. Finally, two clones (no. 22 and 20 of *P. egaliksensis* 1 and 2, respectively) showed a deletion of 34 bp in the ITS1 (e.g. positions 21 and 44 of the ITS alignment, Table 2). From this analysis, we can conclude that most, but not all, ITS sequences of *P. egaliksensis* are homogenised towards the paternal repeat type (*P. nutans*), suggesting that *P. egaliksensis* has progressively lost the maternally inherited ITS sequences. The incomplete homogenisation may indicate that this is still a dynamic process.

Table 2. Clonal variation within *Primula egaliksensis* at the 27 nucleotide positions of the ITS region distinguishing *P. mistassinica* and *P. nutans*. Site numbering follows ITS alignment. Sites identical to the first sequence are represented by dots and gaps by dashes. Site 378 belonging to the 5.8S and separating the ITS1 from ITS2 is italicised.

Taxon	11	21	44	52	57	68	85	90	97	105	106	109	119	126	130	142	143	147	152	153	*178*	184	198	205	254	259	261
P. nutans	C	C	T	T	G	A	A	T	C	T	G	T	T	T	C	T	A	G	T	*C*	A	C	C	C	T	G	C
P. egaliksensis 1																											
c1-13/15/17-18/20-21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
c14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.	.	C	A	.
c16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	*T*	.	.	.	.	.	.	.
c19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
c22	.	–	–	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	.	.	.	.
P. egaliksensis 2																											
c1-15/17-19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
c16	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
c20	.	–	–	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	.	.	.	.
c21	T	T	C	G	.	C	G	C	T	C	.	C	C	C	T	A	G	C	.	.	.	.	.	.	.	.	.
c22	T	T	C	G	.	C	G	C	T	C	.	C	C	C	T	A	G	C	G	.	G	T	.	.	C	A	.
P. mistassinica	T	T	C	G	T	C	G	C	T	C	A	C	C	C	T	A	G	C	G	*T*	G	T	T	T	C	A	T

rDNA loci of P. egaliksensis *have undergone drastic restructuring*

Because our sequencing results suggested that the rDNA of *P. egaliksensis* have undergone concerted evolution, perhaps involving intra- and interlocus recombination, we investigated the number and position of rDNA loci using FISH. Hybridisation with the ITS-PM probe revealed at least 14 signals in *P. mistassinica* (Fig. 2k) and most of them co-localised with the heterochromatic knobs (indicated by arrowheads in Figure 2k). However, only two ITS signals (indicated by arrows in Figure 2k) belonged to the unique pair of 45S rDNA loci detected by FISH (Fig. 2l). Conversely, in *P. nutans*, the ITS-PN probe co-localised with the two 45S rDNA loci (Fig. 2m-n). A similar situation was found in the hybrid *P. egaliksensis*, where the ITS-PE probe co-localised with the two 45S rDNA loci (Fig. 2o-p). Likewise, the parental ITS probes (ITS-PN and ITS-PM) co-localised on two loci of *P. egaliksensis* chromosomes (data not shown). The latter experiment indicate that the parental ITS probes are not different enough (max. 5% divergence) to be distinguished with our experimental conditions (ca. 80% washing stringency) and suggest that *P. egaliksensis* lost the numerous ITS-bearing heterochromatic knobs observed in *P. mistassinica* (see Discussion). Clearly, *P. egaliksensis* has lost one pair of rDNA loci from one parent. To determine the origin of the retained loci, we combined GISH and 45S-FISH and found that the rDNA probe hybridised with chromosomes from *P. nutans* (Figure 2q).

Discussion

Previous studies suggested that *P. egaliksensis* is an allotetraploid originating from hybridisation of the diploids *P. nutans* and *P. mistassinica* (Kelso, 1991; , 1992; Guggisberg *et al.*, unpubl. res.). The present paper unambiguously confirms this hypothesis and provides insights into the evolution of the parental genomes.

Genomic origin and organisation of P. egaliksensis

The chromosome number of *P. egaliksensis* was earlier found to be the sum of the chromosomes found in *P. mistassinica* and *P. nutans* (Kelso, 1991) but the origin of the chromosomes has never been investigated. Using GISH, we can now testify that *P. egaliksensis* possesses 40 chromosomes corresponding to the entire genome complements of *P. mistassinica* and *P. nutans*, thereby confirming its allotetraploid nature.

The hybridisation experiments did not detect major intergenomic rearrangements between the parental genomes (e.g. interchromosomal translocations) in *P. egaliksensis*, suggesting that the chromosomes have retained their parental integrity. Yet, karyological comparisons advocate for intragenomic rearrangements, because the tetraploid lacks the heterochromatic knobs characteristic of its maternal progenitor *P. mistassinica*. Heterochromatic knobs are heavily stained

heterochromatic features, which contain a low density of expressed genes interspersed with numerous retrotransposons and tandemly repeated elements (reviewed by Bennetzen, 2000). The observation that all *P. egaliksensis* chromosomes inherited from *P. mistassinica* do not present these knobs may suggest active mechanisms of chromosome rearrangement near heterochromatic knobs, as shown in maize (Buckler *et al.*, 1999). Alternatively, this result may argue for *P. mistassinica* being polymorphic for heterochromatic knobs, as it is found in maize (Brown, 1949), indicating that the original parent of *P. egaliksensis* may come from a different population of *P. mistassinica* (with no knobs) than the one sampled here (with knobs).

The most compelling evidence of genome restructuring in the allotetraploid *P. egaliksensis* comes from the analysis of rDNA loci. While the sequencing of ITS clones highlighted a vast homogenisation towards the paternal parent *P. nutans*, FISH and GISH analyses demonstrated that the 45S rDNA loci from the maternal parent *P. mistassinica* were lost, in conjunction with the ITS-bearing heterochromatic knobs. The subterminal location of ITS loci in *P. mistassinica* and *P. nutans* may explain the rapid elimination of ITS copies from the maternal parent in the allopolyploid *P. egaliksensis*. Indeed, studies on cotton, tobacco and peonies indicated that loci located near the telomeres are more prone to recombination (leading to possible homogenisation), than sites situated near the centromeres (Wendel *et al.*, 1995; Cronn *et al.*, 1996; Hanson *et al.*, 1996; Zhang and Sang, 1999; Fulnecek *et al.*, 2002). Alternatively, the loss of rDNA genes may be related to nucleolar dominance, a common epigenetic phenomenon in interspecific hybrids whereby only rDNA genes inherited from one parent are transcribed (Reeder, 1985; Pikaard, 2000). Recent investigations on *Brassica* x *Orychophragmus* hybrids suggested that nucleolar dominance may play a role in chromosome stabilisation by inducing genome-specific rearrangements (Li and Ge, 2007). However, remnant nucleotidic signatures of *P. mistassinica* in *P. egaliksensis* ITS clones imply that the loss of maternal ITS sequences occurred later than in the first hybrid generation, since recombination apparently occurred between *P. nutans* and *P. mistassinica* ITS repeats.

Primula egaliksensis *vs. other allopolyploid models*

According to the nucleo-cytoplasmic interaction (NCI) hypothesis proposed by Gill (p. 48, 1991), newly formed polyploids must undergo rapid structural chromosomal changes to lift the 'sterility resulting from the adverse interaction between the male nuclear genome and both the nuclear and cytoplasmic genomes of the female'. Molecular and cytogenetic investigations on synthetic allotetraploid lines of tobacco support this hypothesis, because first-generation polyploids are the genomic sum of their parents, but interchromosomal translocations can be detected after three generations, along with changes in number and composition of rDNA loci (Skalická *et al.*, 2003; Skalická *et al.*, 2005; Lim *et al.*, 2006; Lim *et al.*, 2007).

Primula egaliksensis is an intersectional hybrid (Kelso, 1991, 1992; Richards, 2002), and thus the lack of major intergenomic rearrangements between the parental genomes may be a consequence of the considerable genetic divergence between the putative parents. In synthetic allopolyploids, frequencies of intergenomic recombination were shown to be positively correlated with degrees of divergence between the putative progenitors (Song *et al.*, 1995), but studies on hybrids of distantly related mouse strains also indicated that substantial chromosomal divergence suppresses recombination (Shao *et al.*, 2001). Hence, there might be a positive correlation between the genetic divergence of the parents and the frequency of interchromosomal translocations in the hybrids until this genetic divergence reaches a level where homeologous recombination becomes extremely rare or even impossible.

The NCI hypothesis further predicts that the paternal genome should evolve more rapidly than the maternal one because it functions within an alien maternal genomic environment (Gill, 1991). This assumption has been supported by recent studies on synthetic tobacco (Skalická *et al.*, 2003; Skalická *et al.*, 2005), but comparable assays on natural *Nicotiana* allopolyploids showed that genetic changes may also be targeted at the maternal genome donor (Lim *et al.*, 2000; Kovarík *et al.*, 2004; Clarkson *et al.*, 2005). The present data on the natural allotetraploid *P. egaliksensis* argue for genome restructuring primarily affecting the genome of the maternal progenitor (*P. mistassinica*). Yet, paternally-targeted rearrangements at the gene level cannot be excluded since our investigations were restricted to whole-chromosome structure and rDNA loci.

Primula egaliksensis is supposed to have originated during the Pleistocene glaciations (i.e. between 1.8 million-10 000 years ago) as a result of repeated contact between its putative parents following glacial advancement and retreat (Kelso, 1991, 1992; Richards, 2002; Guggisberg *et al.*, unpubl. res.). The efficiency of *in situ* hybridisation of the parental genomes onto *P. egaliksensis* chromosomes suggests a relatively high degree of conservation of the progenitor sequences in the hybrid and corroborates the hypothesis of a 'geologically young' hybrid with a maximum age of ca. 1 million years. Indeed, recent studies on natural allopolyploids of *Nicotiana* showed that the effectiveness of GISH is considerably reduced after 1 million years of genome evolution and fails after 5 million years of divergence (Clarkson *et al.*, 2005; Lim *et al.*, 2007).

The concomitant loss of ITS-bearing heterochromatic knobs and 45S rDNA sites of *P. mistassinica* in *P. egaliksensis* may be indicative of on-going genome diploidisation (Leitch and Bennett, 2004; Ma and Gustafson, 2005), attesting to the dynamic nature of polyploid taxa (Soltis and Soltis, 1993; Wendel, 2000; Soltis *et al.*, 2003). Future investigations on the *P. egaliksensis* species complex will be aimed at (i) dating the origin of the allopolyploid; (ii) assessing the copy number of 45S rDNA repeats to ascertain that the rDNA loci of *P. egaliksensis* were lost and not

fused and (iii) identifying the ITS repeats borne on heterochromatic knobs of *P. mistassinica*, as they may not be linked to functional 45S rDNA loci (Maggini *et al.*, 1991; Stupar *et al.*, 2002; Lim *et al.*, 2004b).

Acknowledgements

For help with the collection of tissue samples, we thank Bruce Bennett, Guilhem Mansion, Pamela Eveleigh, Wendy Maovlic and all colleagues listed in Table 1. We also acknowledge John Richards for his advice on seed germination, Guilhem Mansion and two anonymous reviewers for comments on earlier versions of the manuscript and Prof. Ingo Schubert for help in the interpretation of heterochromatic knobs. This work was supported by the University of Zurich and by the EPIGENOME Network of Excellence of the European Union to UG. Finally, AG would like to thank the Georges und Antoine Claraz-Schenkung and the Swiss Academy of Sciences for funding the field work.

References

Adams, K. L., and J. F. Wendel. 2005. Polyploidy and genome evolution in plants. Current Opinion in Plant Biology 8:135-141.

Ainouche, A. K., A. Baumel, and A. Salmon. 2004. *Spartina anglica* C. E. Hubbard: a natural model system for analysing early evolutionary changes that affect allopolyploid genomes. Biological Journal of the Linnean Society 82:475-484.

Barker, R. F., N. P. Harberd, M. G. Jarvis, and R. B. Flavell. 1988. Structure and evolution of the intergenic region in a ribosomal DNA repeat unit of wheat. Journal of Molecular Biology 201:1-17.

Baum, D. A., R. L. Small, and J. F. Wendel. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Systematic Biology 47:181-207.

Baumel, A., M. Ainouche, R. Kalendar, and A. H. Schulman. 2002. Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C. E. Hubbard (Poaceae). Molecular Biology and Evolution 19:1218-1227.

Bennetzen, J. L. 2000. The many hues of plant heterochromatin. Genome Biology 1: reviews107.1-107.4.

Bretagnolle, F., F. Felber, F. G. Calame, and P. Küpfer. 1998. La polyploïdie chez les plantes. Botanica Helvetica 108:5-37.

Brown, A. H. D. 1949. Numbers and distribution of chromosome knobs in United States maize. Genetics 34:524-536.

Bruun, H. G. 1932. Cytological studies in *Primula*, with special reference to the relation between the karyology and taxonomy of the genus. Symbolae Botanicae Upsalienses 1:1-239.

Buckler, E. S., IV, T. L. Phelps-Durr, C. S. Keith Buckler, R. K. Dawe, J. F. Doebley, and T. P. Holtsford. 1999. Meiotic drive of chromosomal knobs reshaped the maize genome. Genetics 153:415-426.

Campbell, C. S., M. F. Wojciechowski, B. G. Baldwin, L. A. Alice, and M. J. Donoghue. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Molecular Biology and Evolution 14:81-90.

Chen, Z. J., and Z. Ni. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays 28:240-252.

Clarkson, J. J., K. Y. Lim, A. Kovarík, M. W. Chase, S. Knapp, and A. R. Leitch. 2005. Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae). New Phytologist 168:241-252.

Cronn, R. C., X. Zhao, A. H. Paterson, and J. F. Wendel. 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA diploid and allopolyploid cottons. Journal of Molecular Evolution 42:685-705.

Elder, J. F. J., and B. J. Turner. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. Quarterly Review of Biology 70:297-320.

Fulnecek, J., K. Y. Lim, A. R. Leitch, A. Kovarík, and R. Matyasék. 2002. Evolution and structure of 5S rDNA loci in allotetraploid *Nicotiana tabacum* and its putative parental species. Heredity 88:19-25.

Gill, B. S. 1991. Nucleo-cytoplasmic interaction (NCI) hypothesis of genome evolution and speciation in polyploid plants. Pages 48-53 *in* Nuclear and Organellar Genomes of Wheat Species: Proceedings of Dr. H. Kihara Memorial International Symposium on Cytoplasmic Engineering in Wheat (T. Sasakuma, and T. Kinoshita, eds.). Kihara Memorial Foundation, Yokohama, Japan.

Guggisberg, A., G. Mansion, S. Kelso, and E. Conti. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*. New Phytologist 171:617-632.

Hanson, R. E., M. N. Islam-Faridi, E. A. Percival, C. F. Crane, Y. Ji, T. D. McKnight, D. M. Stelly, and H. J. Price. 1996. Distribution of 5S and 18S-28S rDNA loci in a tetraploid cotton (*Gossypium hirsutum* L.) and its putative diploid ancestors. Chromosoma 105:55-61.

Henry, Y., M. Bedhomme, and G. Blanc. 2006. History, protohistory and prehistory of the *Arabidopsis thaliana* chromosome complement. Trends in Plant Science 11:267-273.

Hultgård, U.-M. 1990. Polyploidy and differentiation in N European populations of *Primula* subgenus *Aleuritia*. Sommerfeltia 11:117-135.

Hultgård, U.-M. 1993. *Primula scandinavica* and *P. stricta* - patterns of distribution, variation, reproductive strategies and migrations. Opera Botanica 121:35-43.

Kelso, S. 1991. Taxonomy of *Primula* sects. *Aleuritia* and *Armerina* in North America. Rhodora 93:67-99.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. Arctic and Alpine Research 24:82-87.

Kovarík, A., R. Matyásek, K. Y. Lim, K. Skalická, B. Koukalová, S. Knapp, M. Chase, and A. R. Leitch. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. Biological Journal of the Linnean Society 82:615-625.

Leitch, I. J., and M. D. Bennett. 2004. Genome downsizing in polyploid plants. Biological Journal of the Linnean Society 82:651-663.

Levin, D. A. 2002. The Role of Chromosomal Change in Plant Evolution. Oxford University Press, New York, NY, USA.

Levy, A. A., and M. Feldman. 2004. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. Biological Journal of the Linnean Society 82:607-613.

Li, Z.-Y., and X.-H. Ge. 2007. Unique chromosome behavior and genetic control in *Brassica* x *Orychophragmus* wide hybrids: a review. Plant Cell Reports 26:701-710.

Lim, K. Y., R. Matyásek, A. Kovarík, and A. R. Leitch. 2004a. Genome evolution in allotetraploid *Nicotiana*. Biological Journal of the Linnean Society 82:599-606.

Lim, K. Y., A. Kovarík, R. Matyásek, M. Bezdek, C. P. Lichtenstein, and A. R. Leitch. 2000. Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units. Chromosoma 109:161-172.

Lim, K. Y., K. Soucková-Skalická, V. Sarasan, J. J. Clarkson, M. W. Chase, A. Kovarík, and A. R. Leitch. 2006. A genetic appraisal of a new synthetic *Nicotiana tabacum* (Solanaceae) and the Kostoff synthetic tobacco. American Journal of Botany 93:875-883.

Lim, K. Y., A. Kovarík, R. Matyásek, M. Chase, J. J. Clarkson, M. A. Grandbastien, and A. R. Leitch. 2007. Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytologist 175:756-763.

Lim, K. Y., K. Skalická, B. Koukalová, R. A. Volkov, R. Matyásek, V. Hemleben, A. R. Leitch, and A. Kovarík. 2004b. Dynamic changes in the distribution of a satellite homologous to intergenic 26-18S rDNA spacer in the evolution of *Nicotiana*. Genetics 166:1935-1946.

Liu, B., and J. F. Wendel. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. Molecular Phylogenetics and Evolution 29:365-379.

Liu, B., C. L. Brubaker, G. Mergeai, R. C. Cronn, and J. F. Wendel. 2001. Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome 44:321-330.

Long, E. O., and I. B. Dawid. 1980. Repeated genes in eukaryotes. Annual Review of Biochemistry 49:727-764.

Lysak, M., P. Fransz, and I. Schubert. 2006. Cytogenetic analyses of *Arabidopsis*. Pages 173-186 *in* Arabidopsis Protocols (J. Salinas, and J. J. Sanchez-Serrano, eds.). Humana Press, Totowa, NJ, USA.

Ma, X.-F., and J. P. Gustafson. 2005. Genome evolution of allopolyploids: a process of cytological and genetic diploidization. Cytogenetics and Genome Research 109:236-249.

Madlung, A., and L. Comai. 2004. The effect of stress on genome regulation and structure. Annals of Botany 94:481-495.

Maggini, F., R. Cremonini, C. Zolfino, G. F. Tucci, R. D'Ovidio, V. Delre, C. DePace, G. T. Scarascia Mugnozza, and P. G. Cionini. 1991. Structure and chromosomal localization of DNA sequences related to ribosomal subrepeats in *Vicia faba*. Chromosoma 100:229-234.

Mason-Gamer, R. 2004. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. Systematic Biology 53:25-37.

Mast, A. R., S. Kelso, and E. Conti. 2006. Are any primroses (*Primula*) primitively monomorphic? New Phytologist 171:605-616.

Mast, A. R., S. Kelso, A. J. Richards, D. J. Lang, D. M. S. Feller, and E. Conti. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. International Journal of Plant Sciences 162:1381-1400.

Matyásek, R., K. Y. Lim, A. Kovarík, and A. R. Leitch. 2003. Ribosomal DNA evolution and gene conversion in *Nicotiana rustica*. Heredity 91:268-275.

McClintock, B. 1984. The significance of responses of the genome to challenge. Science 226:792-801.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Annual Review of Ecology and Systematics 34:401-437.

Pikaard, C. S. 2000. The epigenetics of nucleolar dominance. Trends in Genetics 16:495-500.

Pontes, O., N. Neves, M. Silva, M. S. Lewis, A. Madlung, L. Comai, W. Viegas, and C. S. Pikaard. 2004. Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. Proceedings of the National Academy of Science, USA 101:18240-18245.

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29:467-501.

Reeder, R. H. 1985. Mechanisms of nucleolar dominance in animals and plants. Journal of Cell Biology 101:2013-2016.

Richards, J. 2002. *Primula*. B. T. Batsford, London, UK.

Shao, C., P. J. Stambrook, and J. A. Tischfield. 2001. Mitotic recombination is suppressed by chromosomal divergence in hybrids of distantly related mouse strains. Nature Genetics 28:169-172.

Skalická, K., K. Y. Lim, M. Matzke, A. R. Leitch, and A. Kovarík. 2005. Preferential elimination of repeated DNA sequences from the paternal, *Nicotiana tomentosiformis* genome donor of a synthetic, allotetraploid tobacco. New Phytologist 166:291-303.

Skalická, K., K. Y. Lim, R. Matyásek, B. Koukalová, A. R. Leitch, and A. Kovarík. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. American Journal of Botany 90:988-996.

Soltis, D. E., and P. S. Soltis. 1993. Molecular data and dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243-273.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2003. Advances in the study of polyploidy since *Plant Speciation*. New Phytologist 161:173-191.

Soltis, P. S. 2005. Ancient and recent polyploidy in angiosperms. New Phytologist 166:5-8.

Song, K., P. Lu, K. Tang, and T. C. Osborn. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Science, USA 92:7719-7723.

Stupar, R. M., J. Song, A. L. Tek, Z. Cheng, F. Dong, and J. Jiang. 2002. Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. Genetics 162:1435-1444.

Swofford, D. L. 1999. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), version 4. Sinauer Associates.

Vogelmann, H. W. 1956. A Biosystematic Study of *Primula mistassinica* Michx. University of Michigan, Ann Arbor, MI, USA.

Wagner, A., N. Blackstone, P. Artwright, M. Dick, B. Misof, P. Snow, G. P. Wagner, J. Bartels, M. Murtha, and J. Pendleton. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Systematic Biology 43:250-261.

Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen, J. J. Lee, B. Watson, T. Kagochi, L. Comai, and Z. J. Chen. 2004. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. Genetics 167:1961.

Wendel, J. F. 2000. Genome evolution in polyploids. Plant Molecular Biology 42:225-249.

Wendel, J. F., A. Schnabel, and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proceedings of the National Academy of Science, USA 92:280-284.

Zhang, D., and T. Sang. 1999. Physical mapping of ribosomal RNA genes in peonies (*Paeonia*, Paeoniaceae) by fluorescent *in situ* hybridization: implications for phylogeny and concerted evolution. American Journal of Botany 86:735-740.

Chapter 5

Concluding remarks and future perspectives

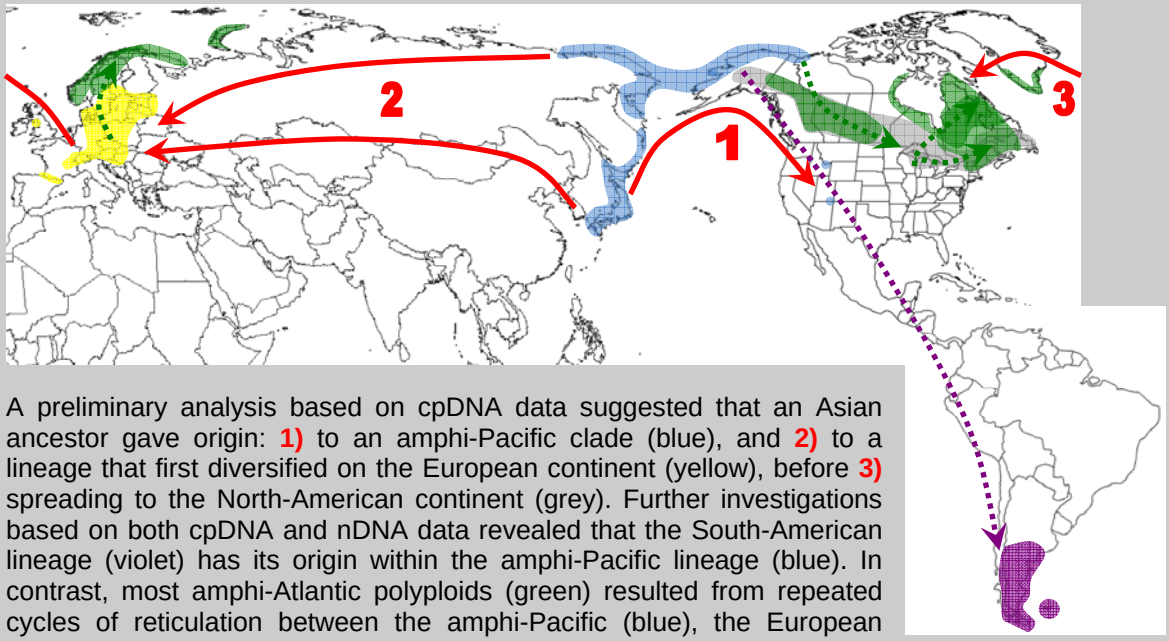
Major findings

The main goal of the present thesis was to disentangle the reticulation patterns in *Primula* sect. *Aleuritia* subsect. *Aleuritia* (hereafter called *Aleuritia*), and to investigate the consequences of polyploidy on the evolution of breeding systems, using an integrative and multifaceted approach. We first investigated the evolutionary history and biogeography of *Aleuritia* within a detailed phylogenetic framework. Preliminary data based on chloroplast DNA (cpDNA) sequences only (cf. **Chapter 2**) confirmed the monophyly of *Aleuritia*, including *P. egaliksensis* (taxonomically ascribed to sect. *Armerina*; Kelso, 1991, 1992), as earlier hypothesised by Mast *et al.* (2001; 2006), and suggested that an Asian ancestor gave origin to an amphi-Pacific clade and to a lineage that diversified on the European and American continents, respectively. This lineage may have spread westward across the central Asian mountain ranges or via the arctic shorelines (see Box 1). Further investigations including both cpDNA and nuclear DNA (nDNA) data (cf. **Chapter 3**) indicated that the South-American octoploid *P. magellanica* originated within the amphi-Pacific lineage, and advocated for repeated cycles of reticulation between the amphi-Pacific, the European and the North-American lineages for most other polyploids within *Aleuritia* (see Box 1).

Indeed, the use of different types of data (i.e. direct DNA sequences from maternally and biparentally inherited genomes, cloned sequences from a nuclear marker, and *in situ* hybridisation techniques; cf. **Chapters 3-4**) and multiple analytical approaches (i.e. generation of cpDNA and nDNA trees and their comparison via consensus networks,

identification of polymorphism and recombination in cloned nDNA sequences), in combination with pre-existing knowledge from morphological, caryological and distributional data, allowed us (i) to infer an allopolyploid origin for most polyploids within *Aleuritia*, including *P. egaliksensis* (hypothesised to be an intersectional hybrid between a member of sect. *Aleuritia* and *Armerina*, respectively; Kelso, 1991, 1992), and (ii) to propose a recurrent origin for at least two of them, *P. laurentiana* and *P. stricta*. Sole the tetraploids *P. borealis* and *P. halleri* and the octoploid *P. magellanica* may have originated via autopolyploidisation.

Box 1: Biogeographic history of *Aleuritia*



A preliminary analysis based on cpDNA data suggested that an Asian ancestor gave origin: **1)** to an amphi-Pacific clade (blue), and **2)** to a lineage that first diversified on the European continent (yellow), before **3)** spreading to the North-American continent (grey). Further investigations based on both cpDNA and nDNA data revealed that the South-American lineage (violet) has its origin within the amphi-Pacific lineage (blue). In contrast, most amphi-Atlantic polyploids (green) resulted from repeated cycles of reticulation between the amphi-Pacific (blue), the European (yellow) and the North-American (grey) lineages.

Our study (cf. **Chapters 2-3**) further corroborated the assumption that widespread, diploid *P. mistassinica*-like and *P. farinosa*-like ancestors played a central role in the evolution of polyploid series on the American and European continents, respectively (Bruun, 1932; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993). However, it also showed that additional diploid lineages, likely related to the extant *P. alcalina*, *P. anvilensis*, *P. modesta*, and *P. specuicola*, were involved in the origin of the American polyploids. These diploids have very narrow distributions and currently occur in areas that were ice-free throughout the Pleistocene (Ono, 1985; Siegert, 2001). Hence, they may represent remnants of a more broadly distributed ancestral diploid that became fragmented during the advancement of Pleistocene glaciers and survived in refugial areas only (Kelso, 1991, 1992).

Altogether, the hybrid make-up of most polyploids, the recurrent formation inferred for some of them, and the involvement of both widespread and narrow endemic diploids in the origin of extant polyploid complexes are compatible with the secondary contact model

(Stebbins, 1984, 1985) earlier invoked to explain speciation in *Aleuritia* (Bruun, 1932; Vogelmann, 1956; Hultgård, 1990; Kelso, 1991, 1992; Hultgård, 1993; Richards, 2002). Preliminary results from *in situ* hybridisation (ISH) experiments on the intersectional hybrid *P. egaliksensis* (cf. **Chapter 4**) further suggest that polyploids may be "geologically young" and likely formed during the last million years.

Our phylogenetic analyses (cf. **Chapter 2**) inferred a diploid, heterostylous most recent common ancestor for *Aleuritia* and multiple origins for the polyploid, homostylous lineages. They further confirmed the proposed correlation between polyploidy and homostyly. We argued for (i) increased rates of recombination (Song *et al.*, 1995; Wendel, 2000; Soltis *et al.*, 2003), and (ii) reduced inbreeding depression (Lande and Schemske, 1985; Thompson and Lumaret, 1992), to explain the frequent breakdown of heterostyly in polyploid taxa. The former hypothesis is corroborated by our ISH experiments (cf. **Chapter 4**), which revealed no large-scale, but drastic small-scale genomic rearrangements following allopolyploidisation. Because most polyploids are of hybrid origin (cf. **Chapters 3-4**), it is though unclear whether hybridisation or polyploidisation was the driving force for the evolution of homostyly. The answer to this question may lie in the genetic make-up of *P. borealis*, the sole polyploid heterostyle belonging to *Aleuritia*. Indeed, our results indicated no evidence of hybridisation in *P. borealis*, and therefore suggested that hybridisation (Kelso, 1992), rather than chromosome doubling, has triggered the shift in breeding system.

Comparisons of distributional ranges between homostylous and heterostylous species (cf. **Chapter 2**) showed that autogamous taxa expanded further north into previously glaciated areas than allogamous ones, implying that the switch to self-compatibility may have represented a selective advantage in the colonisation of empty ecological niches freed up after glacier retreats (Baker, 1955; Stebbins, 1957). Indeed, heterostylous populations of *Aleuritia* growing at the ice-sheet margins may have experienced lower reproductive success due to mate and pollen limitation (Thomson, 1981; Hengeveld and Haeck, 1982; McCall and Primack, 1992; Totland, 1994; Jacquemyn *et al.*, 2002). Under such ecological conditions, selection for reproductive assurance may have favoured the establishment of homostylous, self-compatible mutants that were independent of both mate density and pollinator activity (Stebbins, 1957; Fausto *et al.*, 2001; Kalisz *et al.*, 2004). Overall, the high frequency of polyploid, autogamous species in the Arctic may be the product of selection for increased selfing ability in habitats where pollination is unreliable (Molau, 1993).

Outlook

Biogeography of arctic plants

Few studies have focused on the biogeographic history of widespread, cold-adapted plant genera of the Northern Hemisphere, and even less have incorporated a temporal dimension, i.e. have estimated absolute ages via molecular dating (see Box 2; Scheen *et al.*, 2004; Schneeweiss *et al.*, 2004). The reasons for the presently low number of studies addressing the spatial and temporal origin of the arctic flora are diverse. First, an exhaustive sampling of circumboreal plant groups presents substantial logistic problems, especially in some remote areas such as Siberia, where "missing links" might be located. Second, the taxonomic complexity resulting from frequent cycles of allopolyploidisation (Abbott and Brochmann, 2003; Brochmann *et al.*, 2004) and its direct consequences on phylogenetic reconstructions (e.g. incongruence or low resolution due to conflicting phylogenetic signal; Rieseberg and Ellstrand, 1993; Linder and Rieseberg, 2004; Vriesendorp and Bakker, 2005) may severely complicate broad-scale investigations of arctic plant species complexes. Lastly, the supposedly young age of the arctic flora and resulting lack of fossils for node calibration render molecular dating analyses difficult.

Box 2. Principles of molecular dating*

Molecular dating relies on the assumption that molecular differences accumulate among sequences as time elapses. Hence, temporal landmarks can be incorporated as calibration points into phylogenetic trees, in order to convert relative divergence times into absolute divergence times (e.g. millions of years). Under **rate constancy**, the genetic distances between sequences are proportional to the time elapsed since their divergence (i.e. evolve in a clocklike fashion). When the substitution rate varies over time or across lineages, the system is said to evolve under **rate heterogeneity** (i.e. non-clocklike).

Three sources of information are commonly used for **calibration**: (i) fossils, the most common source; (ii) ages of geological events (e.g. uplift of mountain ranges or opening of land bridges); (iii) age estimates derived from independent molecular dating inferences.

*after Magallón (2004)

The arctic flora probably originated during the Quaternary (1.8 million years ago [mya]–present), because forest grew at high latitudes during most of the Tertiary (65–1.8 mya) and tundra did not develop before the Late Pliocene (3.6–1.8 mya; Matthews Jr. and Ovenden, 1990; Murray, 1995; McIver and Basinger, 1999). It most likely derived from founding stocks located in the high mountain ranges of both central Asia and North-America (Hultén, 1937; Murray, 1995), and achieved a circumboreal distribution by migration across the Bering Land Bridge (BLB), which connected north-east Siberia to north-west North-America following the drop in sea level during the Quaternary glaciations (Elias *et al.*, 1996). Furthermore, recent molecular data (e.g. Abbott *et al.*, 2000; Hagen *et al.*, 2001; and reviewed in Brochmann *et al.*, 2003) suggest that the North Atlantic did not form a strong barrier to dispersal during the

Quaternary, but rather constituted another colonisation route for plants (e.g. via drifting ice or migratory birds; Nordal, 1987). Finally, speciation is thought to have been driven by the cyclic climate changes of the Quaternary, which lead to the fragmentation and differentiation of diploid populations during glacial advancement, and recurrent cycles of allopolyploidisation following secondary contacts after glacial retreat (Stebbins, 1984, 1985; Abbott and Brochmann, 2003; Brochmann *et al.*, 2004).

The aforementioned hypotheses are supported by two molecular studies aimed at reconstructing the organismal diversification of arctic plant species complexes through space and time. Investigations on the arctic-alpine genus *Androsace* s. l. (including the genus *Douglasia* – Primulaceae) suggested a centre of origin in East Asia during the Miocene (23.8–5.3 mya), and advocated for a colonisation of the amphi-Beringian region by European taxa between the Late Pliocene and Pleistocene (i.e. between 3.6–0.01 mya; Schneeweiss *et al.*, 2004). In contrast, phylogenetic investigations on the circumpolar genus *Cerastium* (Caryophyllaceae) indicated an Old World origin and at least two migration events into North-America, one across the BLB during the Miocene and another across the northern Atlantic during the Quaternary (Scheen *et al.*, 2004). The origin of the arctic high-polyploids was estimated to be 1.48–0.44 mya (Scheen *et al.*, 2004), possibly resulting from recurrent episodes of range expansions and contractions during the Quaternary glaciations, as earlier hypothesised by Stebbins (1984; 1985).

To solve the problem raised by the poor fossil record in *Androsace*, Schneeweiss *et al.* (2004) first estimated the temporal origin of *Androsace* on a phylogeny of Primulaceae and related families using the divergence times provided by Wikström *et al.* (2001), and then applied this age estimate to a more detailed phylogeny of *Androsace* and related genera. By contrast, Scheen *et al.* (2004) employed a simple molecular clock assumption (see Box 2) based on the opening of the BLB, to roughly estimate the minimum age of some clades. Both approaches suffer from unavoidable drawbacks, the former from substantial error rates and the latter from the non-independency of the geological calibration (i.e. circularity), but they both provide first insights into the temporal origin of arctic-alpine plants. One way to circumvent the latter bias would be to approximate dates of divergence using different, published rates of evolution under a molecular clock assumption (e.g., Kay *et al.*, 2006), or to estimate divergence times with newly developed coalescence-based methods (e.g., Spellman and Klicka, 2006; Lemmon *et al.*, 2007).

Box 3. Distribution of ice cover (white), tundra (dark grey) and exposed continental shelves (dotted) in the Northern Hemisphere during the Last Glacial Maximum*

During the glacial periods of the Quaternary, most of the Northern Hemisphere was covered by ice and arctic plants had to survive in disjunct, ice-free areas called **refugia**. Evidence from pollen and fossil record indicate that the landmass spanning between north-east Siberia and north-west North-America (known as **Beringia**) represented a major refugium for the arctic flora at that time. This hypothesis has been corroborated by several molecular studies (on the assumption that glacial refugia encompass high levels of genetic diversity), but other ice-free areas may have served as refugia as well, notably (i) the regions south of the ice sheets in both Eurasia and North-America, (ii) several exposed continental shelves in northern Siberia, Scandinavia and eastern North-America, (iii) and nunataks in high-mountain ranges (i.e. ice-free mountain tops).

*after Abbott and Brochmann (2003)

In accordance with findings on *Androsace* and *Cerastium*, *Primula* sect. *Aleuritia* seems to have originated in central Asia, before spreading to Europe and North-America, where recurrent cycles of reticulation resulted in the origin of several polyploid series (cf. Chapters 2-5). Preliminary results from enhanced taxon sampling (including Russian and Caucasian material [notably of subsect. *Algida*] and increased infraspecific variation) actually indicate repeated exchanges between Europe, North-America, and central or north-east Siberia, most likely both across the BLB and the North Atlantic (Guggisberg *et al.*, unpubl. res.). In addition, the extensive infraspecific sampling suggests several glacial refugia (e.g. in Beringia and south of the Great Lakes in North-America; see Box 3) and may allow the identification of potential suture zones (i.e. zones of secondary contacts; Guggisberg *et al.*, unpubl. res.). The nearly perfect additivity observed in *P. stricta* and *P. scandinavica* and the absence of cpDNA haplotype diversity across their whole geographic range further suggest a recent origin for at least some polyploid taxa (Guggisberg *et al.*, unpubl. res.). Finally, preliminary dating analyses based on an ITS phylogeny and assuming rate constancy over time indicate a diversification of diploid lineages between the Late Miocene and Late Pliocene (Guggisberg

et al., unpubl. res.), depending on the substitution rates provided by Kay *et al.* (2006). Yet, no temporal framework has been developed for the entire lineage (including polyploid taxa).

To my knowledge, sole the two papers presented above have tentatively addressed the evolutionary history of high-arctic plant genera in both space and time (Scheen *et al.*, 2004; Schneeweiss *et al.*, 2004). Furthermore, few studies have investigated the genetic composition of arctic plants on a broad, circumboreal scale, in view to identify glacial refugia (Abbott *et al.*, 2000; Alsos *et al.*, 2005; Eidesen *et al.*, 2007b; Eidesen *et al.*, 2007a; Schönswetter *et al.*, 2007). In this sense, future investigations on *Aleuritia* implementing an enhanced infraspecific sampling and a temporal dimension would be a timely contribution to the currently poor knowledge of the biogeography of arctic flora. They would notably allow answering following questions: (i) Did Quaternary glaciations play a central role in the evolution of *Aleuritia*, i.e. are bursts of polyploidisation associated with the climatic oscillations prevailing at that time? Is there a temporal concordance between the origin of European and American lineages? (ii) Does the biogeographic history of *Aleuritia* fit the palaeogeological and palaeoecological data? Does the split between Asian and North-American lineages coincide with the opening of the BLB? Does the split between North- and South-American lineages coincide with the opening of the Panama Isthmus? (iii) Where did *Aleuritia* survive the Pleistocene ice age? Did Beringia constitute a major refugium for *Aleuritia*? Did diploid and polyploid lineages share the same refugia?

Association between polyploidy and self-fertilisation

The switch from an obligate outcrossing to a self-compatible mating system brings in three advantages for plants: (i) selfing, diploid individuals pass on two copies of their genome to their descendants (instead of only one in case of outcrossing; cf. automatic transmission advantage, Fisher, 1941); (ii) self-fertilisation allows the fixation of co-adapted gene complexes (cf. selection for local adaptation, Stebbins, 1957); and (iii) selfing may serve as a reproductive assurance mechanism (Baker, 1955; Stebbins, 1957). The present study on *Aleuritia* favours the latter assumption, because self-compatible homostyles tend to occupy higher latitudes in previously glaciated areas (cf. Chapter 2). While this observation may constitute a plausible explanation for the observed correlation between homostyly and high-latitudinal distribution, it is insufficient to evaluate what caused the simultaneous association with polyploidy in *Aleuritia*. Three reasons (which are not mutually exclusive) have been invoked for the apparent correlation between polyploidy and self-compatibility in angiosperms (Stebbins, 1950; Grant, 1956; Barringer, 2007; Husband *et al.*, 2008): (i) polyploidy causes the disruption of self-incompatibility systems (Barrett, 1988; Richards,

1997); (ii) polyploids are buffered against inbreeding depression (Lande and Schemske, 1985; Husband and Schemske, 1997); and (iii) neopolyploids are more likely to establish if self-compatible (Levin, 1975; Rodríguez, 1996; Ramsey and Schemske, 1998; Rausch and Morgan, 2005).

Previous works on *Primula* have suggested that homostyly results from recombination within the heterostyly linkage group (Dowrick, 1956; Wedderburn and Richards, 1992). We have argued that most polyploids within *Aleuritia* switched to a homostylous breeding system, because polyploids usually undergo major genomic rearrangements after their formation (cf. Chapter 2; Wendel, 2000; Adams and Wendel, 2005; Chen and Ni, 2006). The occurrence of a tetraploid heterostyle in *Aleuritia* (*P. borealis*) however raised the question, whether polyploidisation indeed triggered the switch in breeding system. For the degree of intergenomic recombination may be lower in auto- vs. allopolyploids (Song *et al.*, 1995; Shao *et al.*, 2001), we have argued that the maintenance of heterostyly in *P. borealis* might be the consequence of an autopolyploid origin of this taxon (cf. Chapter 2). In other words, hybridisation, rather than polyploidisation, may be the driving force for the switch to homostyly in *Aleuritia*. This assumption could be verified by exploring the genomic composition of *P. borealis* in greater detail. Similarly, the investigation of species of sects. *Parryi* and *Auricula* would be of highest interest, because these sections are exclusively composed of heterostylous polyploids (Richards, 2002).

Stebbins (1957; 1971) early observed that autopolyploids tend to be less self-compatible than allopolyploids (reviewed in Husband *et al.*, 2008). He notably argued that autopolyploids do not tolerate selfing, because they would quickly become homozygous (Stebbins, 1957). Since then, three theoretical models have been proposed to assess the fitness costs of selfing, specifically inbreeding depression (ID), in organisms characterised by disomic vs. polysomic inheritance. According to the overdominance model (deleterious alleles are treated as dominant), autopolyploids suffer greater ID than their diploid progenitors, because allelic interactions decreases more rapidly upon selfing than in diploids (Busbice and Wilsie, 1966). By contrast, the partial dominance model of Lande and Schemske (1985; deleterious alleles are treated as partially recessive) suggests that tetraploidy reduces ID, because in tetraploids homozygosity evolves more slowly upon selfing than in diploids. Finally, under the partial dominance model refined by Ronfort (1999; deleterious alleles possess different dominance coefficients), ID in autopolyploids may either increase or decrease with selfing rate, depending on the level of dominance of the deleterious alleles. ID in allopolyploids has not

been addressed explicitly, probably because the inheritance of allopolyploids is thought to be disomic as in diploids (reviewed in Ramsey and Schemske, 2002).

The first empirical studies on autopolyploids corroborated the predictions of the partial dominance model developed by Lande and Schemske (1985), for they reported a lower ID in tetraploid than in diploid populations of *Chamerion* (Onagraceae; Husband and Schemske, 1997) and *Anthericum* (Liliaceae; Rosquist, 2001), respectively. More recent investigations on *Chamerion angustifolium* however indicated that the magnitude of ID may increase between newly-formed and well-established autopolyploids, implying 'that selfing may be ephemeral and that selection ultimately favours mixed or outcrossed mating systems in autopolyploids' (p. 195, Husband *et al.*, 2008). By contrast, experiments on *Amsinckia* (Boraginaceae) revealed that homostylous allotetraploids presented higher rates of ID than heterostylous, diploid ones (Johnston and Shoen, 1996). These surprising results must though be treated cautiously, because the populations were shown to be highly selfing, meaning that lethal or sublethal mutations may not constitute an important source of ID anymore (Johnston and Shoen, 1996). Overall, these preliminary observations clearly emphasise the need for better empirical knowledge on ID in polyploid plants, and denote the importance of comparing auto- and allopolyploids with their diploid progenitors at different evolutionary stages.

Finally, the different selfing rates noticed between auto- and allopolyploid plants (Stebbins, 1957, 1971; Husband *et al.*, 2008) may reflect the mode of formation of the respective types of polyploids. On the one hand, Gustafsson (1948) stated that allopolyploids are less likely to form in self-incompatible than in outcrossing species. On the other hand, Grant (1956) and Stebbins (1950; 1957) suggested that selfing is particularly beneficial to the stabilisation of hybrids, because it increases the likelihood of chromosome doubling (i.e. allopolyploidy) through the union of unreduced gametes. While all theoretical models show that selfing favours the establishment of neopolyploids in diploid populations (cf. minority cytotype exclusions principle; Levin, 1975; Rodríguez, 1996; Ramsey and Schemske, 1998; Rausch and Morgan, 2005), Ramsey and Schemske (1998) predicted different rates of auto- and allopolyploid formation upon selfing and outcrossing. According to their calculations (which did not take the frequency of F_1 hybrid formation into account), allopolyploids form more often than autopolyploids, because hybrids produce more diplogametes and therefore form many allopolyploids spontaneously during F_1 self crosses (Ramsey and Schemske, 1998). By contrast, allopolyploids probably only form upon outcrossing 'after a breakdown of reproductive isolation, as might be observed in disturbed habitats' (p. 495, Ramsey and

Schemske, 1998). Unfortunately, these estimates were based on very few data taken from agricultural and horticultural systems, once again highlighting the need for additional, empirical studies on natural, polyploid complexes looking at the pathways and rates of polyploid formation.

To sum up, the difficulty in interpreting the observed correlation between polyploidy and self-compatibility in *Aleuritia* may result from the mutual effects of (1) polyploidy on the costs of selfing and (2) selfing on the likelihood of polyploid formation and establishment, and calls attention to additional, integrative investigations on natural, polyploid systems aimed at answering following questions: (i) Does recombination constitute the genetic mechanism that causes the disruption of the heterostyly syndrome? If not, is it possible that other consequences of polyploidisation (e.g. activation of transposable elements) lead to the switch in breeding systems, as suggested by a recent study on the *S*-locus (Manfield *et al.*, 2005)? (ii) What is the rate of auto- vs. allopolyploid formation in both selfing and outcrossing species? And what are the genetic and phenotypic characteristics limiting their respective establishment in diploid populations? (iii) Do auto- and allopolyploids preferentially form through the union of unreduced gametes? If so, do auto- and allopolyploids form via different pathways (cf. uni- vs. bilateral pathway)? (iv) How will diploid-polyploid contact zones evolve in light of upcoming, climatic and environmental changes (e.g. global warming, anthropogenic disturbance) with respect to cytotype frequencies and proportions of homostyles vs. heterostyles? Will further hybridisation constitute an escape from extinction, i.e. facilitate adaptive radiation (see Seehausen, 2004; Willis *et al.*, 2006)? May future changes, e.g. in pollinator fauna, cause the disappearance of diploid, heterostylous populations, and favour the establishment of homostylous, polyploid mutants (see Washitani, 1996)?

References

Abbott, R. J., and C. Brochmann. 2003. History and evolution of the arctic flora: in the footsteps of Eric Hultén. Molecular Ecology 12:299-313.

Abbott, R. J., L. C. Smith, R. I. Milne, R. M. M. Crawford, K. Wolff, and J. Balfour. 2000. Molecular analysis of plant migration and refugia in the Arctic. Science 289:1343-1346.

Adams, K. L., and J. F. Wendel. 2005. Polyploidy and genome evolution in plants. Current Opinion in Plant Biology 8:135-141.

Alsos, I. G., T. Enkelskjøn, L. Gielly, P. Taberlet, and C. Brochmann. 2005. Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. Molecular Ecology 14:2739-2753.

Baker, H. G. 1955. Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347-349.

Barrett, S. C. H. 1988. The evolution, maintenance, and loss of self-incompatibility systems. Pages 98-124 *in* Plant Reproductive Ecology: Patterns and Strategies (J. Lovett Doust, and L. Lovett Doust, eds.). Oxford University Press, New York, NY, USA.

Barringer, B. C. 2007. Polyploidy and self-fertilization in flowering plants. American Journal of Botany 94:1527-1533.

Brochmann, C., T. M. Gabrielsen, I. Nordal, J. Y. Landvik, and R. Elven. 2003. Glacial survival or *tabula rasa*? The history of North Atlantic biota revisited. Taxon 52:417-450.

Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A.-C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. Biological Journal of the Linnean Society 82:521-536.

Bruun, H. G. 1932. Cytological studies in *Primula*, with special reference to the relation between the karyology and taxonomy of the genus. Symbolae Botanicae Upsalienses 1:1-239.

Busbice, T. H., and C. P. Wilsie. 1966. Inbreeding depression and heterosis application to *Medicago sativa* L. Euphytica 15:52-67.

Chen, Z. J., and Z. Ni. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays 28:240-252.

Dowrick, V. P. J. 1956. Heterostyly and homostyly in *Primula obconica*. Heredity 10:219-236.

Eidesen, P. B., T. Carlsen, U. Molau, and C. Brochmann. 2007a. Repeatedly out of Beringia: *Cassiope tetragona* embraces the Arctic. Journal of Biogeography 34:1559-1574.

Eidesen, P. B., I. G. Alsos, M. Popp, Ø. Stensrud, J. Suda, and C. Brochmann. 2007b. Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species. Molecular Ecology 16:3902-3925.

Elias, S. A., S. K. Short, C. H. Nelson, and H. H. Birks. 1996. Life and times of the Bering land bridge. Nature 382:60-63.

Fausto, J. A., V. M. Eckhart, and M. A. Geber. 2001. Reproductive assurance and the evolutionary ecology of self-pollination in *Clarkia xantiana* (Onagraceae). American Journal of Botany 88:1794-1800.

Fisher, R. A. 1941. Average excess and average effect of a gene. Annals of Eugenics 11:53-63.

Grant, V. 1956. The influence of breeding habit on the outcome of natural hybridization in plants. The American Naturalist 90:319-322.

Gustafsson, A. 1948. Polyploidy, life-form and vegetative reproduction. Hereditas 34:1-22.

Hagen, A. R., H. Giese, and C. Brochmann. 2001. Trans-Atlantic dispersal and phylogeography of *Cerastium arcticum* (Caryophyllaceae) inferred from RAPD and SCAR markers. American Journal of Botany 88:103-112.

Hengeveld, R., and J. Haeck. 1982. The distribution of abundance. I. Measurements. Journal of Biogeography 9:303-316.

Hultén, E. 1937. Outline of the History of Arctic and Boreal Biota during the Quaternary Period: their Evolution during and after the Glacial Period as Indicated by the Equiformal Progressive Areas of Present Plant Species. J. Cramer, Lehre, Germany.

Hultgård, U.-M. 1990. Polyploidy and differentiation in N European populations of *Primula* subgenus *Aleuritia*. Sommerfeltia 11:117-135.

Hultgård, U.-M. 1993. *Primula scandinavica* and *P. stricta* - patterns of distribution, variation, reproductive strategies and migrations. Opera Botanica 121:35-43.

Husband, B. C., and D. W. Schemske. 1997. The effect of inbreeding in diploid and tetraploid populations of *Epilobium angustifolium* (Onagraceae): implications for the genetic basis of inbreeding depression. Evolution 51:737-746.

Husband, B. C., B. Ozimer, S. L. Martin, and L. Pollock. 2008. Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. International Journal of Plant Sciences 169:195-206.

Jacquemyn, H., R. Brys, and M. Hermy. 2002. Patch occupancy, population size and reproductive success of a forest herb (*Primula elatior*) in a fragmented landscape. Oecologia 130:617-625.

Johnston, M. O., and D. J. Shoen. 1996. Correlated evolution of self-fertilization and inbreeding depression: an experimental study of nine populations of *Amsinckia* (Boraginaceae). Evolution 50:1478-1491.

Kalisz, S., D. W. Vogler, and K. M. Hanley. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. Nature 430:884-887.

Kay, K. M., J. B. Whittall, and S. A. Hodges. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. BMC Evolutionary Biology 6.

Kelso, S. 1991. Taxonomy of *Primula* sects. *Aleuritia* and *Armerina* in North America. Rhodora 93:67-99.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. Arctic and Alpine Research 24:82-87.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24-40.

Lemmon, E. M., A. R. Lemmon, and D. C. Cannatella. 2007. Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudarcis*). Evolution 61:2086-2103.

Levin, D. A. 1975. Minority cytotype exclusion in local plant populations. Taxon 24:35-43.

Linder, C. R., and L. H. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. American Journal of Botany 91:1700-1708.

Magallón, S. A. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. International Journal of Plant Sciences 165:S7-S21.

Manfield, I. W., V. K. Pavlov, J. Li, H. E. Cook, F. Hummel, and P. M. Gilmartin. 2005. Molecular characterization of DNA sequences from the *Primula vulgaris S*-locus. Journal of Experimental Botany 56:1177-1188.

Mast, A. R., S. Kelso, and E. Conti. 2006. Are any primroses (*Primula*) primitively monomorphic? New Phytologist 171:605-616.

Mast, A. R., S. Kelso, A. J. Richards, D. J. Lang, D. M. S. Feller, and E. Conti. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. International Journal of Plant Sciences 162:1381-1400.

Matthews Jr., J. V., and L. E. Ovenden. 1990. Late Tertiary plant macrofossils from localities in arctic/subarctic North America - a review of the data. Arctic 43:364-392.

McCall, C., and R. B. Primack. 1992. Influence of flower characteristics, weather, time of day, and season on insect visitation rates in three plant communities. American Journal of Botany 79:434-442.

McIver, E. E., and J. F. Basinger. 1999. Early Tertiary floral evolution in the Canadian High Arctic. Annals of the Missouri Botanical Garden 86:523-545.

Molau, U. 1993. Relationships between flowering phenology and life history strategies in tundra plants. Arctic and Alpine Research 25:391-402.

Murray, D. F. 1995. Causes of arctic plant diversity: origin and evolution. Pages 21-32 *in* Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences (F. S. Chapin III, and C. Körner, eds.). Springer Verlag, Heidelberg, Germany.

Nordal, I. 1987. Tabula rasa after all? Botanical evidence for ice-free refugia in Scandinavia reviewed. Journal of Biogeography 14:377-388.

Ono, Y. 1985. Late glacial paleoclimate reconstructed from glacial and periglacial landforms in Japan. Geographical Review of Japan, Series B 57:80-100 (in Japanese).

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29:467-501.

Ramsey, J., and D. W. Schemske. 2002. Neopolyploidy in flowering plants. Annual Review of Ecology and Systematics 33:589-639.

Rausch, J. H., and M. T. Morgan. 2005. The effect of self-fertilization, inbreeding depression, and population size on autopolyploid establishment. Evolution 59:1867-1875.

Richards, A. J. 1997. Plant Breeding Systems, 2nd edition. Cambridge University Press, Cambridge, UK.

Richards, J. 2002. *Primula*. B. T. Batsford, London, UK.

Rieseberg, L. H., and N. C. Ellstrand. 1993. What can molecular and morphological markers tell us about plant hybridization? Critical Review in Plant Sciences 12:213-241.

Rodríguez, D. J. 1996. A model for the establishment of polyploidy in plants. The American Naturalist 147:33-46.

Ronfort, J. 1999. The mutation load under tetrasomic inheritance and its consequences for the evolution of the selfing rate in autotetraploid species. Genetical Research 74:31-42.

Rosquist, G. 2001. Reproductive biology in diploid *Anthericum ramosum* and tetraploid *A. liliago* (Anthericaceae). Oikos 92:143-152.

Scheen, A.-C., C. Brochmann, A. K. Brysting, R. Elven, A. Morris, D. E. Soltis, P. S. Soltis, and V. A. Albert. 2004. Northern hemisphere biogeography of *Cerastium* (Caryophyllaceae): insights from phylogenetic analysis of noncoding plastid nucleotide sequences. American Journal of Botany 91:943-952.

Schneeweiss, G. M., P. Schönswetter, S. Kelso, and H. Niklfeld. 2004. Complex biogeographic patterns in *Androsace* (Primulaceae) and related genera: evidence from phylogenetic analyses of nuclear internal transcribed spacer and plastid *trn*L-F sequences. Systematic Biology 53:856-876.

Schönswetter, P., J. Suda, M. Popp, H. Weiss-Schneeweiss, and C. Brochmann. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. Molecular Phylogenetics and Evolution 42:92-103.

Seehausen, O. 2004. Hybridization and adaptive radiation. Trends in Ecology and Evolution 19:198-207.

Shao, C., P. J. Stambrook, and J. A. Tischfield. 2001. Mitotic recombination is suppressed by chromosomal divergence in hybrids of distantly related mouse strains. Nature Genetics 28:169-172.

Siegert, M. J. 2001. Ice Sheets and Late Quaternary Environmental Change. Wiley, Chichester, UK.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2003. Advances in the study of polyploidy since *Plant Speciation*. New Phytologist 161:173-191.

Song, K., P. Lu, K. Tang, and T. C. Osborn. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Science, USA 92:7719-7723.

Spellman, G. M., and J. Klicka. 2006. Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). Proceedings of the Royal Society of London, Series B: Biological Sciences 273:3057-3063.

Stebbins, G. L. 1950. Variation and Evolution in Plants. Columbia University Press, New York, NY, USA.

Stebbins, G. L. 1957. Self-fertilization and population variability in the higher plants. The American Naturalist 91:337-354.

Stebbins, G. L. 1971. Chromosomal Evolution in Higher Plants. Edward Arnold, London, UK.

Stebbins, G. L. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. Botanica Helvetica 94:1-13.

Stebbins, G. L. 1985. Polyploidy, hybridization, and the invasion of new habitats. Annals of the Missouri Botanical Garden 72:824-832.

Thompson, J. D., and R. Lumaret. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. Trends in Ecology and Evolution 7:302-307.

Thomson, J. D. 1981. Spatial and temporal components of resource assessment by flower-feeding insects. Journal of Animal Ecology 50:49-59.

Totland, Ø. 1994. Influence of climate, time of day and season, and flower density on insect flower visitation in alpine Norway. Arctic and Alpine Research 26:66-71.

Vogelmann, H. W. 1956. A Biosystematic Study of *Primula mistassinica* Michx. University of Michigan, Ann Arbor, MI, USA.

Vriesendorp, B., and F. T. Bakker. 2005. Reconstructing patterns of reticulate evolution in angiosperms: what can we do? Taxon 54:593-604.

Washitani, I. 1996. Predicted genetic consequences of strong fertility selection due to pollinator loss in an isolated population of *Primula sieboldii*. Conservation Biology 10:59-64.

Wedderburn, F., and A. J. Richards. 1992. Secondary homostyly in *Primula* L.: evidence for the model of the 'S' supergene. New Phytologist 121:649-653.

Wendel, J. F. 2000. Genome evolution in polyploids. Plant Molecular Biology 42:225-249.

Wikström, N., V. Savolainen, and M. W. Chase. 2001. Evolution of the angiosperms: calibrating the family tree. Proceedings of the Royal Society of London, Series B: Biological Sciences 268:2211-2220.

Willis, B. L., M. J. H. van Oppen, D. J. Miller, S. V. Vollmer, and D. J. Ayre. 2006. The role of hybridization in the evolution of reef corals. Annual Review of Ecology, Evolution and Systematics 37:489-517.

Acknowledgements

I would like to thank…/Ich danke…/Je remercie…

…**Guilhem Mansion**, qui m'a fidèlement accompagnée tout au long de ce « périple », pour son inlassable soutien, son amour et son inépuisable patience; grâce à lui, un petit « Maulwurf » curieux et attentif sommeillera toujours en moi.

…meinen Eltern **Marlise und Ulrich Guggisberg**, sowie meiner Schwester **Sarah Guggisberg-Rüesch** und ihrer Familie für ihre treue Unterstützung, ihre Liebe und ihre Geduld, in guten wie in schlechten Zeiten; ohne sie hätte ich dieses Projekt nie zu Ende gebracht.

…my supervisor **Elena Conti** for giving me the opportunity to broaden my knowledge in plant evolution; her manifold interests provided the setting for many inspiring discussions.

…**Célia Baroux**, **Sylvia Kelso** and **Ueli Grossniklaus** for their insightful contributions; thanks to Célia *in situ* hybridisations are not a mystery for me anymore!

…dem Promotionskomitee und der Jury bestehend aus **Peter Linder**, **Florian Schiestl**, **Jakob Schneller** und **Alex Widmer** für ihren Zeitaufwand und ihre aufschlussreichen Bemerkungen.

…the referees **Brian Husband**, **Andrew Leitch** and **Gonzalo Nieto-Feliner** for judging the present thesis.

…unserer Institutssekretärin **Corinne Burlet** für ihre grenzenlose Hilfsbereitschaft in administrativen und organisatorischen Fragen.

…**Daniel Heinzmann**, **Barbara Keller**, **Sandro Wagen** and **Shuqing Xu** for their precious help and support in the lab.

…**Niklaus Müller** für seinen Einsatz in Sachen Computerprobleme.

…den Bibliothekarinnen **Elena Benetti** und **Barbara Seitz** für ihre Hilfe während der andauernden Literatursuche.

…**Rolf Rutishauser** fürs Korrigieren der deutschen Texte.

…the **Georges und Antoine Claraz-Schenkung** and the **Swiss Academy of Sciences** for supporting my various field trips.

…all colleagues and friends who helped me compiling the plant material necessary to undertake the present study, in particular **Torbjørn Alm**, **Craig Anderson**, **Alan Batten**, **Bruce Bennett**, **Sean Blaney**, **Christian Brochmann**, **Luc Brouillet**, **Edward Buyarski**, **Tor Arne Carlsen**, **Matthew Carlson**, **Mary Beth Cook**, **Reidar Elven**, **Pamela Eveleigh**, **Bruce Ford**, **Jason Grant** and his parents **Wyan and Henry Grant**, **Pekka Halonen**, **Oyumaa Jamsran**, **Emmet Judziewicz**, **Tupuna Kovanen**, **Hörður Kristinsson**, **Wendy Maovlic**, **Michael Oldham**, **Carolyn Parker**, **John Richards**, **Carl Roland**, **Michal Ronikier**, **Peter Schönswetter**, **Ayako Shimono**, **Heidi Solstad**, **Andreas Tribsch**, **Ida Trift** and **Didrik Vanhoenacker**.

…the following herbaria for providing invaluable information on plant localities: ACAD, ALA, ALAJ, ALTA, AMNH, CAFB, CAN, CDFN, DAO, GB, HAM, ICEL, LKHD, MT, NFLD, NHIC, NSPM, O, QFA, S, SASK, SCFQ, SLU, SMI, TRH, TROM, TRTE, S, UBC, UME, UNB, UQTR, UPS, UWPG, WAT, WIN, WLK.

…all colleagues and friends of the Institute of Systematic Botany, who rendered my stay so lovely, especially **Karina Arroyo**, **Julia Born**, **Cyril Guibert**, **Zsofià Hock**, **Urs Landergott**, **Brigitte Marazzi**, and **Péter Szövenyi**.

…everybody I shared the office with, namely **Michael Pirie**, **Danielle Sombra-Stäheli**, **Beatrix Itten**, **Catarina Rydin**, **Severin Dietschi** and **Philip Taxboeck**, for such a friendly working atmosphere.

Curriculum vitae

Last name	GUGGISBERG
First name	Alessia
Date of birth	October 14, 1978
Nationality	Swiss from Zimmerwald/BE

Education

2003-2008 PhD thesis at the Institute of Systematic Botany, University of Zurich, Switzerland, under the supervision of Prof. Elena Conti "Polyploidy and evolution of breeding systems in an arctic-alpine species complex of *Primula*: an integrative phylogenetic and cytogenetic study"

2001-2002 MsC thesis at the Evolutionary Botany Laboratory, University of Neuchâtel, Switzerland, under the supervision of Prof. François Bretagnolle and Dr. Guilhem Mansion "Origin of *Centaurium bianoris* (Gentianaceae), an allotetraploid species endemic to Majorca (Balearic Islands/Spain)"

1997-2001 undergraduate studies in Biology at the University of Neuchâtel, Switzerland

1994-1997 Matura Typus C at the "Lycée Denis-de-Rougemont", Neuchâtel, Switzerland

Teaching experience

2003-2008 teaching assistant at the Institute of Systematic Botany, University of Zurich, Switzerland

2000-2001 volunteer lab assistant at the Institute of Evolutionary Botany, University of Neuchâtel, Switzerland

Other activities

2007 co-organiser of the international congress "Origin and Evolution of Biota in Mediterranean Climate Zones: an Integrative Vision", Zurich, Switzerland

2004-2005 co-supervision of K. Arroyo's MsC thesis "Life history, distribution and genetic diversity of the rare species *Saxifraga florulenta* Moretti, a narrow endemic of the Maritime Alps", Institute of Systematic Botany, University of Zurich, Switzerland

2001 BirdLife Switzerland trainee, Centre Nature ASPO de la Sauge, Cudrefin, Switzerland

1996-present collaborator of the Swiss Ornithological Society

referree for *Australian Systematic Botany*, *Molecular Ecology*, *New Phytologist*, and *Plant Systematics and Evolution*

Publications

Guggisberg, **A**., F. Bretagnolle, and G. Mansion. 2006. Allopolyploid origin of the Mediterranean endemic, *Centaurium bianoris* (Gentianaceae), inferred by molecular markers. Systematic Botany 31: 368-379.

Guggisberg, **A**., G. Mansion, S. Kelso, and E. Conti. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*. New Phytologist 171:617-632.

Guggisberg, **A**., C. Baroux, U. Grossniklaus, and E. Conti. 2008. Genomic origin and organisation of the allopolyploid *Primula egaliksensis* investigated by *in situ* hybridisations. Annals of Botany doi:10.1093/aob/mcn026.

Guggisberg, **A**., G. Mansion, and E. Conti. tentatively accepted. Disentangling reticulate evolution in an arctic-alpine polyploid complex. Systematic Biology.

Szövényi, P., K. Arroyo, **A**. **Guggisberg**, and E. Conti. submitted. Effects of Pleistocene glaciations on the genetic structure of *Saxifraga florulenta* (Saxifragaceae), a rare endemic of the Maritime Alps. Taxon.

Mansion, G., F. Selvi, **A. Guggisberg**, and E. Conti. submitted. Origin of Mediterranean insular endemics in the Boraginaceae: integrative evidence from molecular dating and ancestral area reconstruction. Journal of Biogeography.

Guggisberg, **A**., G. Mansion, S. Kelso, and E. Conti. in preparation. Systematics and biogeography of *Primula* sect. *Aleuritia*: insights from enhanced taxon sampling.

Conferences

Guggisberg, **A.**, G. Mansion, C. Baroux, S. Kelso, U. Grossniklaus, and E. Conti. 2008. When does hybridisation lead to the formation of new species? Answers and perspectives from *Primula* sect. *Aleuritia*. Biology08, Annual Meeting of the Swiss Zoological, Botanical and Mycological Societies, Lausanne, Switzerland. (poster)

Guggisberg, **A.**, G. Mansion, C. Baroux, S. Kelso, U. Grossniklaus, and E. Conti. 2008. When does hybridisation lead to the formation of new species? Answers and perspectives from *Primula* sect. *Aleuritia*. NCCR Plant Survival International Conference on "Plant Species Concept and Evolution", Neuchâtel, Switzerland. (poster)

Mansion, G., **A. Guggisberg**, F. Selvi, and E. Conti. 2008. Origin of insular, Mediterranean endemics in the Boraginaceae – integrative evidence from molecular dating and ancestral area reconstruction. NCCR Plant Survival International Conference on "Plant Species Concept and Evolution", Neuchâtel, Switzerland. (poster)

Guggisberg, **A.**, G. Mansion, and E. Conti. 2007. Disentangling reticulate evolution: examples from *Primula*. Symposium of the Zurich-Basel Plant Science Centre, Zurich, Switzerland. (poster)

Guggisberg, **A.**, C. Baroux, U. Grossniklaus, and E. Conti. 2007. Genomic origin and organisation of the allopolyploid *Primula egaliksensis* investigated by *in situ* hybridisation. Symposium of the Zurich-Basel Plant Science Centre, Zurich, Switzerland. (poster)

Guggisberg, **A.**, G. Mansion, and E. Conti. 2007. Disentangling reticulate evolution: examples from *Primula*. International congress on "Origin and Evolution of Biota in Mediterranean Climate Zones: an Integrative Vision", Zurich, Switzerland. (poster)

Guggisberg, **A.**, and G. Mansion. 2007. Extensive hybridization within the Mediterranenan *Centaurium* Hill (Gentianaceae). International congress on "Origin and Evolution of Biota in Mediterranean Climate Zones: an Integrative Vision", Zurich, Switzerland. (poster)

Guggisberg, **A.**, G. Mansion, and E. Conti. 2007. Disentangling reticulation patterns: an example from the *Aleuritia*-polyploid complex (*Primula* L.). Biology07, Annual Meeting of the Swiss Zoological, Botanical and Mycological Societies, Zurich, Switzerland. (talk)

Guggisberg, **A.**, G. Mansion, S. Kelso, and E. Conti. 2006. Evolution of biogeographic patterns, ploidy levels and breeding systems in a diploid-polyploid species complex of *Primula*. Botany06, Annual Meeting of the Botanical Society of America, Chico, CA, USA. (poster)

Guggisberg, **A.**, G. Mansion, and E. Conti. 2006. Disentangling reticulate evolution in the *Aleuritia* polyploid complex (*Primula* L.). Evolution06, Stony Brook, NY, USA. (talk)

Guggisberg, **A.**, F. Bretagnolle, and G. Mansion. 2006. Allotetraploid origin of an endemic species of Majorca (*Centaurium bianoris* – Gentianaceae) inferred with molecular markers. Evolution06, Stony Brook, NY, USA. (poster)

Guggisberg, **A.**, S. Kelso, and E. Conti. 2005. Evolution of breeding systems and ploidy levels in *Primula* sect. *Aleuritia*, a paradigm of the secondary contact model. Symposium of the Zurich-Basel Plant Science Centre, Zurich, Switzerland. (poster)

Guggisberg, **A.**, S. Kelso, and E. Conti. 2005. Evolution of breeding systems and ploidy levels in *Primula* sect. *Aleuritia*, a paradigm of the secondary contact model. XVII International Botanical Congress, Vienna, Austria. (invited talk)

Guggisberg, **A.**, S. Kelso, and E. Conti. 2005. Evolution of breeding systems and ploidy levels in *Primula* sect. *Aleuritia*, a paradigm of the secondary contact model. Evolution05, Fairbanks, AK, USA. (poster)

Guggisberg, **A.**, S. Kelso, and E. Conti. 2005. Evolution of breeding systems and ploidy levels in *Primula* sect. *Aleuritia*, a paradigm of the secondary contact model. Biology05, Annual Meeting of the Swiss Zoological, Botanical and Mycological Societies, Basel, Switzerland. (talk)

Guggisberg, **A.**, G. Mansion, and F. Bretagnolle. 2004. Allotetraploid origin of an endemic species to Majorca/Balearic Islands (*Centaurium bianoris* – Gentianaceae) inferred with molecular markers. IX IOPB Meeting, Valencia, Spain. (poster)

Guggisberg, **A.**, A. Mast, S. Kelso, and E. Conti. 2003. Polyploidy, biogeography and evolution of breeding systems in *Primula* sect. *Aleuritia* with special emphasis on an allotetraploid complex: an integrated phylogenetic, population genetic, and molecular cytogenetic approach. Symposium of the Zurich-Basel Plant Science Centre, Zurich, Switzerland. (poster)

Guggisberg, **A.**, A. Mast, S. Kelso, and E. Conti. 2003. Polyploidy, biogeography and evolution of breeding systems in *Primula* sect. *Aleuritia* with special emphasis on an allotetraploid complex: an integrated phylogenetic, population genetic, and molecular cytogenetic approach. Botany03, Annual Meeting of the Botanical Society of America, Mobile, AL, USA. (poster)

Guggisberg, **A.**, G. Mansion, and F. Bretagnolle. 2003. Allotetraploid origin of an endemic species to Majorca/Balearic Islands (*Centaurium bianoris* – Gentianaceae) inferred with molecular markers. International Polyploidy Conference, London, UK. (poster)

Guggisberg, **A.**, G. Mansion, and F. Bretagnolle. 2003. Allotetraploid origin of an endemic species to Majorca/Balearic Islands (*Centaurium bianoris* – Gentianaceae) inferred with molecular markers. Biology03, Annual Meeting of the Swiss Zoological, Botanical and Mycological Societies, Zurich, Switzerland. (talk)

Guggisberg, **A.**, G. Mansion, F. Felber, and F. Bretagnolle. 2002. Origin of *Centaurium bianoris* (Gentianaceae), an allotetraploid species endemic to Majorca (Balearic Islands, Spain). Ecological Genetics Group meeting, Newcastle, UK. (poster)

Grants

2007 900CHF from the Claraz-Schenkung for the visit of Prof. Loren Rieseberg in Vancouver, Canada

2005 4,500CHF from the Claraz-Schenkung for a field trip to Scandinavia

2004 4,700CHF from the Claraz-Schenkung for a field trip to Canada

2003 3,500CHF from the Swiss Academy of Sciences for a field trip to Alaska